

JPRS-UCC-90-003

7 MAY 1990



FOREIGN
BROADCAST
INFORMATION
SERVICE

JPRS Report

Science & Technology

USSR: Computers

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

19981217 116

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

Reproduced From
Best Available Copy

DTIC QUALITY INSPECTED 3

JPRS-UCC-90-003
7 MAY 1990

SCIENCE & TECHNOLOGY

USSR: COMPUTERS

CONTENTS

HARDWARE

Optoelectronic Digital Computer System in Residue Arithmetic for Image Processing [A. A. Akayev, S. Z. Dordoyev; AVTOMETRIYA, No 3, May-Jun 89].....	1
Optoelectronic Signal Processors [M. A. Gofman, Ye. S. Nezhevenko, et al.; AVTOMETRIYA, No 3, May-Jun 89].....	9
Microelectronic Optical Digital Computer Devices [V. M. Yegorov, E. G. Kostsov; AVTOMETRIYA, No 3, May-Jun 89].....	21
Architecture of an Information System Based on a Large- Capacity Holographic Memory [B. V. Vanyushev, N. N. Vyyukhina, et al.; AVTOMETRIYA, No 3, May-Jun 89].....	32
High-Speed Digital Data Memory on an Optical Disk Pack [Yu. V. Vovk, L. V. Vydrin, et al.; AVTOMETRIYA, No 3, May-Jun 89].....	44

Bubble Memories
[N. B. Malinovskiy; UPRAVLYAYUSHCHIYE SISTEMY I MASHINY,
No 5, Sep-Oct 89]..... 62

Multiprocessor Computing System for Simulation of Radio
Systems
[Ye. V. Voronov, A. A. Grigoryev, et al.;
UPRAVLYAYUSHCHIYE SISTEMY I MASHINY, No 5, Sep-Oct 89].... 72

APPLICATIONS

A Parallel Bolder Algorithm for Rotation Operations and Its
Use in Computer Graphics
[Ye. I. Artamonov, Sh.-M. A. Ismailov, et al.;
UPRAVLYAYUSHCHIYE SISTEMY I MASHINY, No 1, Jan 90]..... 85

UDC 681.323:535

Optoelectronic Digital Computer System in Residue Arithmetic
for Image Processing

907G0029A Novosibirsk AVTOMETRIYA in Russian No 3 May-June 1989 (manuscript received 22 Nov
88) pp 48-53

[Article by A.A. Akayev and S.Z. Dordoyev, Frunze: "Optoelectronic Digital Computer System in Residue Arithmetic for Image Processing," published under the rubric "Organization and Structure of Optoelectronic Computer Systems"]

[Text] Introduction

The development of prototype specimens of optical and integral-optical switch elements with acceptable characteristics made it possible to create optoelectronic digital computer systems [OTsVS]. OTsVS are expected to achieve a high performance level due to fast switching elements and the advantages of optics: natural parallelism and easy establishment of connections. In the area of general-purpose computers, where an excessively high operation speed is not necessary, OTsVS are not likely to compete with conventional electronic computer technology, especially with the promise of a new generation of electronic transistors, switching in the picosecond range. On the other hand, in digital image processing, where multidimensional signals have to be processed with a discretization frequency of hundreds of megahertz in real time, electronics has encountered serious difficulties. Hardware implementation of important algorithms of digital signal processing is needed, which requires parallel structures with multiple retunable connections. Structures of this kind can be obtained in specialized optoelectronic computer systems. The present paper suggests an architecture of image processing OTsVS with residue arithmetic, utilizing integral-optical switches as the component base.

Integral-Optical Switches

Integral-optical switches are the closest to commercial introduction among nonlinear optical logic elements. We will describe an electrooptical

waveguide switch or a directed coupler. Two parallel titanium lightguides are formed by diffusion on a lithium niobate (LiNbO_3) backing. Electrodes are then precipitated on the surface. In the absence of a voltage on the electrodes, the energy entered into one waveguide is transmitted to the other waveguide with a high ratio. If a voltage is applied to the electrodes, light flows through the same waveguide without switching. According to estimates in [1], the switching time can be 25-30 ps, with scattering energy of 30 pJ. There are reports of experimental switching frequencies above 12 GHz at low excitation power of 1.5 mW/GHz [2]. The switching energy in that case is 1.5 pJ, which is not achievable in the near future for electronic transistors. French investigators succeeded in reducing the loss introduced by the switch to 0.1-0.2 dB/cm [3]. One of the outputs of the directed coupler is electrical. By installing a photodetector at that input one can have a switch where all inputs and outputs are optical.

Due to the relatively weak electrooptical effect in lithium niobate crystal, the minimum admissible switch length is a few millimeters, which prevents building of compact optoelectronic integrated circuits. In semiconductor structures with multiple quantum wells — GaInAsP/InP and AlGaAs/GaAs — an electrooptical effect greater by several orders of magnitude than in LiNbO_3 , has been obtained. This quantum electrooptical effect, called the Stark effect, is utilized in compact optoelectronic switches on crossing waveguides. The active segment is $\sim 10 \mu\text{m}$. The theoretically feasible waveguide switching angle has been estimated at 10° . The switching time is the relaxation time of the electric dipole, estimated at 0.1 ps. Since quantum phenomena in superlattices consume very little energy (on the order of picojoules), these superlattice switches are a promising component base for OTsVS [2].

Residue Arithmetic

The choice and analysis of number system is an important aspect in developing a computer architecture. Early on in the development of digital optical computation technology, the system of residue classes [SOK] was described as promising [4]. Recently, interest in this nonposition system has reemerged. The computational process in SOK has two advantages, which fit well the specifics of optical data processing: an absence of interbit carries, which is attractive for parallel systems, and the shortness of residues representing the number making it possible to use table arithmetic. Addition and multiplication can be implemented in SOK easily, but division and comparison are virtually impossible. SOK can be an effective number system for a specialized OTsVS with digital image processing, where multiplication and addition are the main operations.

The bases in SOK are n coprime integers m_1, m_2, \dots, m_n , called modules. Any integer A is represented by a set of residues (r_1, r_2, \dots, r_n) , where $r_i = [A]_{m_i}$ is

the residue of division of A by m_i . This representation is unique in a dynamic range of

$$M = \prod_{i=1}^n m_i$$

Suppose that in this system two numbers have been specified: $X = (x_1, x_2, \dots, x_n)$ and $Y = (y_1, y_2, \dots, y_n)$. The arithmetic operation on these numbers is executed as follows: $Z = X * Y = (z_1, z_2, \dots, z_n)$, where $z_i = [(x_i * y_i)]m_i$ ($i = 1, n$). The symbol * stands for an arithmetic operation of addition, subtraction or multiplication. The operations are executed in each SOK module in parallel and independent of other modules.

Several methods have been suggested for executing arithmetic in SOK. In the final analysis, all are based on the second advantage: operating with short numbers and calculating with the aid of tables. For instance, addition and multiplication modulo 5 can be performed of these tables:

	0	1	2	3	4		0	1	2	3	4
0	0	1	2	3	4	0	0	0	0	0	0
1	1	2	3	4	0	1	0	1	2	3	4
2	2	3	4	0	1	2	0	2	4	1	3
3	3	4	0	1	2	3	0	3	1	4	2
4	4	0	1	2	3	4	0	4	3	2	1

Addition
Multiplication

These tables can easily be implemented as arrays of optical logic elements, photodetectors, and integral-optical switches using the method of spatial mapping [1].

Applications in Digital Image Processing

Specialized residue arithmetic OTsVS can be employed effectively for multiplying a matrix by a matrix or a vector by a matrix. Since the latter operation is a subset of the former, it is sufficient to implement multiplication of a vector by a matrix:

$$c_k = \sum_{n=0}^{N-1} a_n b_{nk}, \quad k = \overline{0, N-1},$$

where a_n is an element of the input vector; b_{nk} is an element of the transformation matrix; c_k is an element of the resultant vector. Important algorithms of digital signal processing are reducible to this operation.

Circular convolution of two periodic sequences

$$y_k = \sum_{n=0}^{N-1} x_n h_{k-n}, \quad k = \overline{0, N-1}$$

is reducible to multiplying the input vector X by a Toeplitz matrix H whose elements $h_{kn} = h_{k-n}$. (A Toeplitz matrix is one whose elements on each diagonal are identical, i.e., the elements of neighboring rows are the same, with a shift by one position.)

The correlation of two periodic sequences

$$z_k = \sum_{n=0}^{N-1} x_n h_{n-k}, \quad k = \overline{0, N-1}$$

is also reducible to multiplying the input vector X by a Toeplitz matrix H :

$$Z = X \times H.$$

A digital filter is implemented by the algorithm

$$y_n = -\sum_{j=1}^{M-1} a_j y_{n-j} + \sum_{l=0}^{N-1} b_l x_{n-l}$$

which is the sum of two convolutions.

A retunable connection block, where each input can be linked to each output, is implemented as multiplication of the input vector by a switching matrix made up of zeros and ones. The ones represent connection and the zeros the absence of connection.

An optoelectronic multiplier of a vector by a matrix can find broad applications in digital processing of multidimensional signals.

OTsVS Architecture

The synthesis of a computer system starts from listing requirements for its characteristics, proceeding from an available or planned component base and an application area. As demonstrated above, there are promising optoelectronic switching components with acceptable characteristics. There is also a class of tasks that can be executed effectively by OTsVS. These are functions in digital processing of multidimensional signals on the basis of multiplication of a vector by a matrix. Accordingly, the main requirement for an OTsVS is a high performance output that would enable it to process signals in real time.

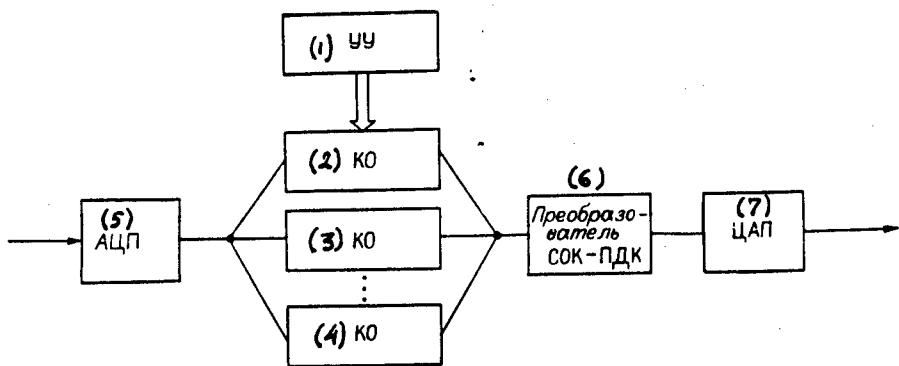


Figure 1. Structure of OTsVS in SOK: ATsP = analog-to-digital converter; KO = channel of processing modulo m_i ; TsAP = digital-to-analog converter; UU = control device.

Key: 1. — UU; 2 — KO; 3 — KO; 4 — KO; 5 — ATsP; 6 — SOK-PDK converter; 7 — TsAP.

Data parallelism is natural and typical for this class of problems. With a view to a full utilization of the advantages of optical systems, OTsVS should be constructed as systems of the OKMD class (single-instruction multiple data stream).

Since a specialized OTsVS is designed to process images in real time, a fully parallel architecture with no addressing is preferable [5]. Only intermediate memory element arrays are needed at the input and output of the processing channel to store information during one cycle. For such matrices one can use PVMS [not further identified] or arrays of nonlinear optical logic elements. A conventional sequential architecture with addressing is also possible. The effect would probably then be accomplished by reducing the execution time of arithmetic operations by a large factor. Future research will help identify the best alternatives.

Figure 1 is an illustration of OTsVS structure with residue arithmetic. An input image is fed to an ATsP [analog-to-digital converter]. A digitized two-dimensional signal is then processed in n independent channels (where n is the number of SOK modules). From channel outputs data are sent to the input of the SOK converter — binary position code [PDK] — and further to TsAP [digital-to-analog converter] for output of the processed image. The system is controlled by a microcomputer, which can also be used for communication with the operator.

Two-dimensional input data in binary code are a sequence of bit planes which in each processing channel is converted to a matrix of space-coded residues (fig. 2).

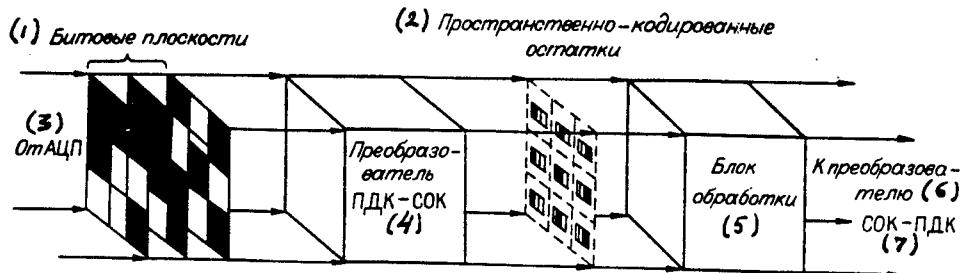


Figure 2. Channel of processing modulo m_i .

Key: 1 — bit planes; 2 — space-encoded residues; 3 — from ATsP; 4 — PDK-SOK converter; 5 — processing block; 6 — to converter; 7 — SOK-PDK.

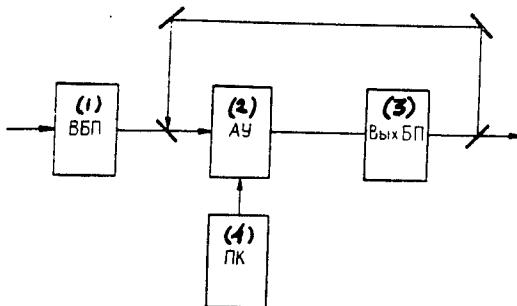


Figure 3. Block of processing modulo m_i ; VBP = input buffer memory; AU = arithmetic device; OutBP = output buffer memory; PK = coefficients memory.

Key: 1 — VBP; 2 — AU; 3 — OutBP; 4 — PK.

Essentially, the entire computational process takes place in the processing block (fig. 3). The arithmetic device executes the operation of multiplication of matrix by matrix or vector by matrix. Addition and multiplication are executed on the basis of matrices of integral-optical switch elements by the table method. To synchronize the operation of the processing block and the converters, PVMS are placed at the input and output as buffer memory. If necessary, an output signal can be forwarded through the optical system to the arithmetic device input. The coefficients of the transformation matrix can be stored in an associative holographic memory.

Main Features of OTsVS

We will estimate the system's speed. We assume that arithmetic operations are carried out in a multifunctional computer module constructed on the basis of a matrix of electrooptical waveguide switches [1]. The light propagation time in this module is approximately 40 ps; the complete cycle of execution of an operation is 3 ns. The time of execution of a matrix operation

consists of the multiplication time and the SOK-PDK and reverse transformation times. If the system executes the multiplication $C = A \cdot B$, where **A** is an $m \times n$ matrix and **B** is an $n \times p$ matrix, then the execution time of this operation is estimated as $m(3 + 0.04n)$ ns. For 512×512 matrices this time is equal to $12 \mu s$. The converter to SOK from a binary code, based on the same universal module, transforms one bit plane in 4 ns. If an array of 32-bit words is received at the input, the PDK-SOK conversion will take 128 ns. The reverse conversion from SOK to the binary code is more difficult. It goes through an intermediate transformation to a system with a mixed base. According to estimates, the time of SOK-PEK conversion is approximately $(200N)$ ns, where N is the dimension of the input data array. If $N = 512$, SOK-PDK conversion takes $102.4 \mu s$. Accordingly, multiplying two 512×512 matrices in OTsVS will take approximately $116 \mu s$.

Take the characteristics of a T9506 Toshiba single-chip VLSI processor [6]. A parallel image processing system based on this VLSI multiplies 512×512 matrices in 26 ms, or by two orders of magnitude slower than a residue arithmetic OTsVS.

Computation accuracy in a computer is measured by the number of bits in a computer word and the form of data representation. For a given number of bits in fixed-point format, a higher accuracy is achieved than with a floating-point format. In the T9506 processor, a 32-bit data format with a fixed point is implemented that allows representing the range of absolute decimal numbers approximately from 1 to 10^9 . For a range of SOK numbers to be 10^9 , 10 first prime numbers can be taken as the system modules:

$$m_{i=1,10} = (2, 3, 5, 7, 11, 13, 17, 19, 23, 29).$$

This means that for the accuracy achieved in a 32-bit fixed-point format it is sufficient for the OTsVS to have 10 independent processing channels with SOK modules of smallest magnitude.

Conclusions

Recent developments in integral-optical switches suggest that they can be viewed as a promising component base for OTsVS. There is a class of digital image processing problems where OTsVS could be widely used. In this context, an architecture of a specialized optoelectronic digital computer system is suggested. The design principles of OTsVS and the data representation in the system are analyzed, and the class of computer systems is chosen. For the number system an nonposition system of residue classes is considered. A structure of image processing OTsVS with residue arithmetic is proposed. The main characteristics of the system — operation speed and precision — are evaluated.

Bibliography

1. Akayev, A.A., and Mayorov, S.A., OPTICHESKIYE METODY OBRABOTKI INFORMATSII [Optical Methods of Information Processing], Moscow, Vysshaya Shkola, 1988.
2. Sumatsu, Y. and Arai, S., "Integral-Optical Approach to Development of Promising Semiconductor Lasers" (Russian translation) IEEE TRANSACTIONS, Vol 75, No 11, 1987.
3. Gallagher, R.T., "Optoelectronic Research in France" (Russian translation), ELEKTRONIKA, Vol 59, No 2, 1986.
4. Orlov, L.A., and Popov, Yu.M., "Optoelectronic Arithmetic Device in a System of Residue Classes," AVTOMETRIYA, No 6, 1972.
5. Huang, A., "Architecture of an Optical Digital Computer" (Russian translation), IEEE TRANSACTIONS, Vol 72, No 7, 1984.
6. Sugai, M., Kanuma, A., Suzuki, K., and Kubo, M., "VLSI-Processor for Image Processing" (Russian translation), IEEE TRANSACTIONS, Vol 75, No 9, 1987.

COPYRIGHT: "Nauka," "Avtometriya," 1989

UDC 681.325.5:535

Optoelectronic Signal Processors

907G0029B Novosibirsk AVTOMETRIYA in Russian No 3 May-June 1989 (manuscript received 22 Dec 88) pp 53-60

[Article by M.A. Gofman, Ye.S. Nezhevenko and V.I. Feldbush (Novosibirsk): "Optoelectronic Signal Processors," published under the rubric "Organization and Structure of Optoelectronic Computer Systems"]

[Text] Optical information processing has been around for quite a while — nearly a quarter of a century. Yet, practical applications of devices based on optical processing are still a rarity. There are several reasons for this: successes in electronic computer technology that have created strong competition for optical technology; the scarcity of a component base; the unpreparedness of the industry for introduction of optical computers. Besides these problems, there are basic shortcomings in optical processing that, in our opinion, have slowed down its development. One flaw is the narrow specialization of optical computers and their lack of flexibility. One has to develop a new device for virtually every new task, involving a large expenditure of time and resources. In electronic computer technology this problem has been solved by general-purpose computers, although recently there has been a switch from general use of these computers to more specialized devices that provide higher speed at a relatively small cost. One such class is digital signal processors [TsSP].

Despite the limited set of basic operations, these processors implement a majority of image and signal processing algorithms; the operation speed of TsSP is significantly higher than with general-purpose computers. This is accomplished by means of hardware implementation of the most time-consuming procedures, a special memory organization and a wide use of pipeline and systolic architectures. When programmed with PPZU [programmable read-only memories] TsSP virtually become specialized devices.

This leads to the question of whether an optical device similar to TsSP could be created, i.e., a device that would on one hand have a universal basic structure but, on the other, would become a specialized device as a result of

programming. In this paper we demonstrate that such computer devices can be created; they are called optoelectronic signal processors [OSP] [1].

The main units of TsSP are [2]: an arithmetic logic device, a multiplier, a device of microprogramming sequence connected to the memory and a controller. OSP generally reproduce TsSP structure. At first glance, this can only be accomplished by constructing OSP according to the principles of digital technology. However, considering the specifics of most signal (image) processing algorithms, it is possible to build OSP on elements of analog optoelectronics. Indeed, multiplication and summation in optics at the analog level is not difficult. Furthermore, these operations can be conducted through multiple channels, which will of course speed up the processor.

Among the basic problems with optical processors are their static quality, the difficulty of modifying a data processing algorithm, and, in a large sense, the impossibility of reprogramming the processor. The problem can be resolved by the following techniques: 1) entering input information as a time sequence of vectors rather than parallel arrays (matrices), as is usual in optical processing; 2) pipelining computational processes; 3) reindexing intermediate and resulting data arrays by a spatial shift (it will be recalled that in optics the space coordinate can be one of the independent variables); 4) readdressing light flows in the course of computations.

The first of these techniques is accomplished either by a natural time modification of signals or by scanning images (such as on television). The second and third techniques are accomplished by dynamic optoelectronic elements, such as acoustooptical PZS [charge-coupled device]-arrays, a row and an array of light emitters with shift registers, etc. The hardware for readdressing of light beams can be based on lightguide, holographic, lens-raster and other optical elements.

According to these suggestions, we will describe a general structure of an optoelectronic signal processor (fig. 1). A memory of input data is needed when the pace of arrival of input data is different from the processing speed. The memory should be constructed of electronic components. It is followed by a digital-to-analog input section. It forms the control pulse for the light emitter, providing a proportionality of the light energy to the input signal value (or the signal component determined by the bit representation of the signal).

In a different version, data represented in an appropriate wordlength format can be entered into OSP as a spatial distribution of density, a refractive index or any other parameter of the modulating medium, which when illuminated by the light emitter produces a spatial distribution of light proportional to the input data. The spatial optical signal is sent to an index decrement block, which will be discussed later. The first block of static readdressing

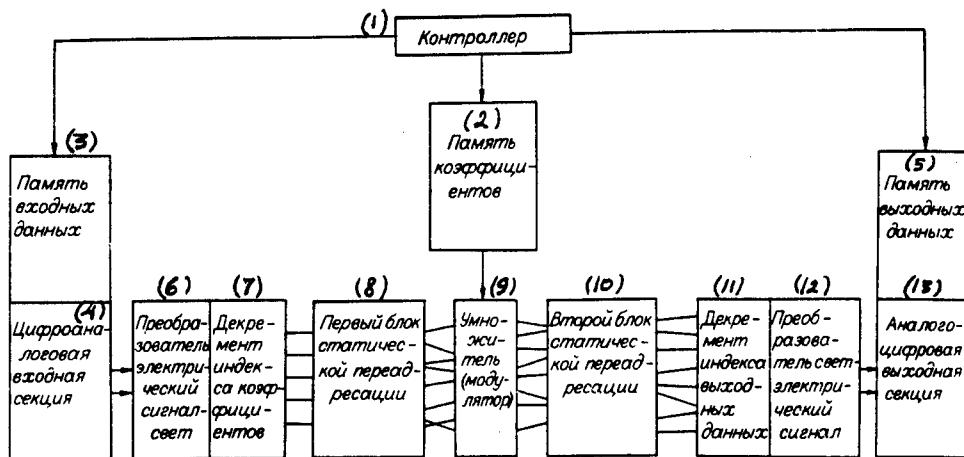


Figure 1.

Key: 1 — controller; 2 — coefficients memory; 3 — input data memory; 4 — digital-to-analog input section; 5 — output data memory; 6 — electric signal-light converter; 7 — decrement of coefficients index; 8 — first static readdressing block; 9 — multiplier (modulator); 10 — second static readdressing block; 11 — output data index decrement; 12 — light-to-electric signal converter; 13 — analog-to-digital output section.

sends the input light flows to appropriate modulator regions. This includes multiplication of input data by coefficients stored in the OSP memory and registered as a spatial distribution of transmittance of the modulator, which is proportional to the coefficients. The indices of the factors are determined by the spatial representation of the input data and the coefficients. When the relative position of the input array and the field of coefficients changes, a reindexing of coefficients (decrement or increment) takes place which is necessary for implementing all kinds of processing algorithms. Hardware implementations of the coefficients index decrement block can be diverse: an acoustooptical modulator, shift registers, etc.

From the modulator the light flows reach the second readdressing block, which operates jointly with the block of input data index decrement. The latter should be typically based on PZS-devices; it executes the following functions: a) it forms the charge relief proportional to the light flows falling upon a photosensitive surface, b) it shifts this relief at the pace of data reindexing and c) it sums up the light flows as the charge relief is being shifted.

The second readdressing block sums up the light flows according to the algorithm and sends them to the specified regions of the photosensitive medium of the PZS-matrix. The combination of data index decrement operations and static readdressing of data makes it possible not only to execute virtually any readdressing of data files, but also to process data in a pipeline mode.

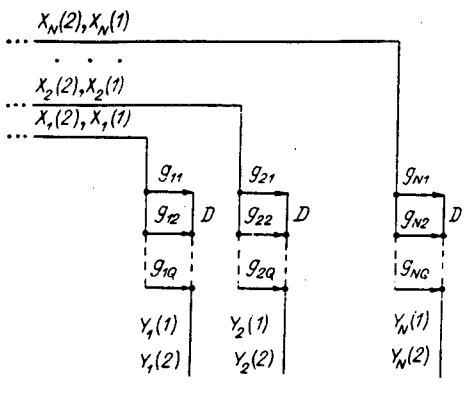


Figure 2.

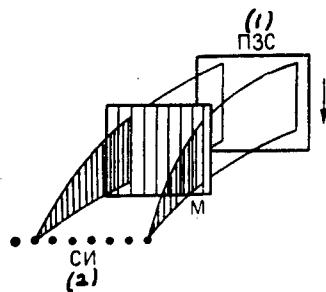


Figure 3.

Key: 1 — PZS; 2 — SI.

We will describe real possibilities of tuning (programming) OSP to execute various specific functions. We will first note the following. Signal processing algorithms can conveniently be represented as signal flow graphs [GPS] [3]. Notations in GPS will be explained during the course of the description of the first example of OSP programming: tuning the system for signal processing by filters with finite-impulse characteristics [KIKh-filter].

§1. Tuning OSP for Multichannel KIKh-Filtration

This operation is described by

$$Y(k) = \sum_{l=1}^L x_n(k-l)g_{nl}, \quad n = 1, 2, \dots, N.$$

GPS implementing this algorithm is shown in fig. 2. The vector processed is $X(m) = [x_1(m), x_2(m) \dots x_N(m)]$, where n is the index of the vector component and m is the time index. A black dot in GPS is an operational graph node, which denotes an arithmetic operation. The graph edge, marked by D ($2D$, $3D$, ...), represents a signal delay by one (two, three, ...) cycles. The edge labeled by a lower-case letter (a , b , g , etc.) denotes multiplication into the respective constant.

The format of GPS for KIKh-filtration points up immediately certain specific features of OSP tuning for execution of this task. Obviously, processing channels are independent (there are no crosslinks); the input data indices are decremented by formation of vectors: time cuts of signals in N channels. The edges D are situated after the edges of multiplication into coefficients

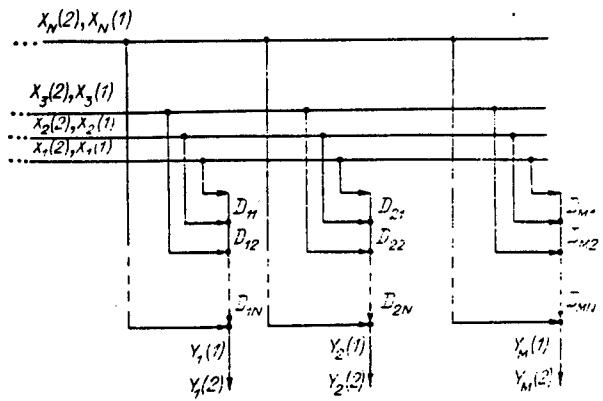


Figure 4.

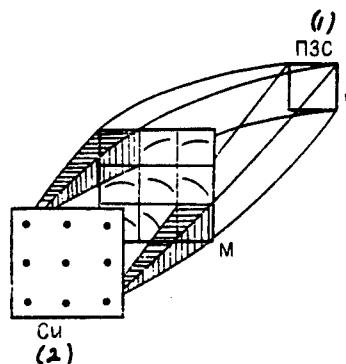


Figure 5.

Key: 1 — PZS; 2 — SI

g. Thus, there is no decrement of coefficient indices; only output data indices are decremented.

A scheme of multichannel KIKh-filtration in OSP is shown in fig. 3. The light emitter [SI] format is a line. The first readdressing block translates a luminescent point to a line. The kernels of transformations g_{n1} are written in the modulator (M). The second readdressing block performs projection (point-to-point). The output data are decremented by means of the output block, built as a PZS-array in the mode of time delay with accumulation [VZN].

§2. Tuning OSP to Process Signals Transmitted by Waves Propagating in Space

This form of data processing is common in seismology, geophysics, radiolocation, hydroacoustics, etc. [4]. The objective is to isolate signal components spreading in a certain direction (this is sometimes referred to as ray formation). If the waves are detected by an array of receivers, an elementary method of ray formation is weighted addition with delay. It is defined by the expression

$$Y_m(k) = \sum_{n=1}^N x_n(k - \tau_{mn}(k)),$$

where $(x_1 \dots x_N)$ is the vector of signal values received by an array of N detectors; k is the time index; m is the ray number.

The graph of the signal flow representing this task is shown in fig. 4, where $D_{mn} = D(\tau_{mn} - \tau_{m,n-1})$, $\tau_{m,0} = 0$. An inspection of this GPS suggests the

following: all input channels take part in formation of each ray; the information from each channel is a function of two arguments: the serial number of the output channel (ray) and the delay $\tau_{mn}(k)$; and the input data indices are decremented by evolution of the input signal in time; output data indices must be decremented.

One possible scheme for tuning OSP to execute this function is presented in fig. 5. The format of light emitters is a matrix. The energy of each light emitter is translated onto a portion of the modulator, which is a two-dimensional region with coordinates m, k . There are N regions, i.e., the first readdressing block is a "point-two-dimensional region" unit. Each region on the modulator is a mask that forms a light flow as a line; the configuration of the line depends on the position of the sensors that generate the signals processed. All the regions are then combined to one region on the photosensitive layer of the PZS-array responsible for the decrement of the index k , i.e., second readdressing block is "N regions — one region." The resulting signal is represented as an $m \times k$ array, where m is the ray number and k is the time index.

§3. Tuning OSP for Multichannel Spectral Data Analysis

Suppose that we want to carry out a spectral analysis of data received simultaneously in N channels, i.e., to compute the Fourier coefficients:

$$Y_{kn} = \sum_{m=1}^M x_n(m)a_k(m),$$

where

$$\begin{aligned} a_k(m) &= \sin mk/M, \quad k = 1, 3, 5, \dots, N-1; \quad a_k(m) = \cos mk/M, \\ &\quad k = 0, 2, 4, \dots, N. \end{aligned}$$

Of the possible variants of spectral analysis, we take the one that matches the OSP philosophy and is determined by the set of basic operations. The GPS for this variant is shown in fig. 6.

We will list its main features: the input data indices are decremented as a result of evolution of processes over time; the edge of multiplication into a constant is situated in the GPS between two edges with index D (therefore, both indices of coefficients and of output data must be decremented); and, the columns of output data correspond to different spectral components (the coefficients of sine and cosine decomposition); the rows correspond to different channels.

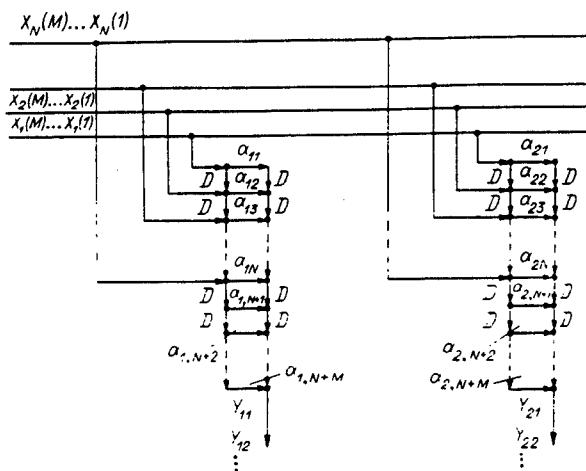


Figure 6.

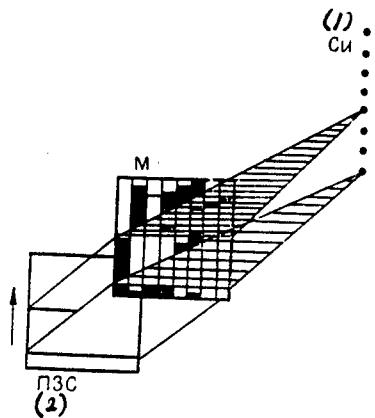


Figure 7.
Key: 1 — SI; 2 — PZS;

At first glance, it seems that all channels take part in forming the Fourier expansion in each GPS channel. However, that is not the case: since the indices are decremented both for coefficients and for output data, the input and output flows move synchronously, without "mixing," while coefficients remain in place. This produces the effect of multiplying the signal by a moving sine or cosine component.

The OSP scheme for spectral analysis is shown in fig. 7. The source of emission is a line of light emitters oriented in the direction of register shift in the PZS-array. The energy of each light emitter is transformed to a line perpendicular to the orientation of the range of emitters, i.e., the first readdressing block is "point-line." During a signal discretization interval, that will be called a cycle, N light emitters are activated (N is the number of channels), corresponding to the signal time cut. After the end of the cycle the next time cut is shifted one position along the range of light emitters which corresponds to the decrement of coefficients. At the same time, the bit map of the PZS-array is shifted in the same direction.

The modulator is a mask with transmittance functions in the direction of the charge shift that are proportional to $(1 + \sin k/N)$ and $(1 + \cos k/N)$. The second readdressing block provides projection of the light distribution from the modulator onto the photosensitive surface of the PZS-array (the "point-point" addressing).

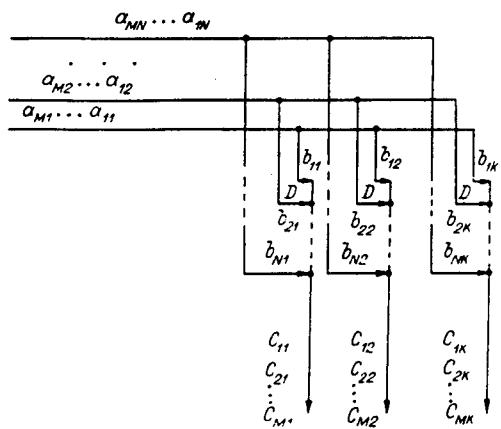


Figure 8.

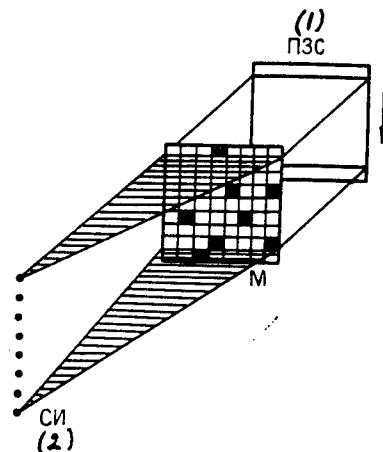


Figure 9.

Key: 1 — PZS; 2 — SI.

§4. Tuning OSP for Matrix Multiplication

Suppose that we want to find the result of multiplication of two matrices $C = A \times B$, i.e., an element of the matrix

$$C_{mk} = \sum_{n=1}^N a_{mn} b_{nk}$$

The GPS for matrix multiplication in OSP philosophy can be illustrated by fig. 8. The main features of the graph are the following:

- one of the matrices (matrix A in the figure) is entered into OSP with a shift of rows, i.e., each next row is entered into the processor with a shift by one element relative to the preceding row; as a result, a rectangular matrix becomes rhomboid;
- there is no decrementing of coefficients; only the output data indices are decremented synchronously with the entry of input data;
- the output data format is a rectangular matrix that is fed out of OSP row by row at the pace of the input of the multiplicand matrix with a delay determined by the size of the multiplicand and multiplier matrices (pipeline delay).

The tuning of OSP for GPS of matrix multiplication is illustrated by fig. 9. The input block is a line range of light emitters. The first readdressing block transforms the light emitter energy to a line ("point-line" addressing). The matrix B is recorded on the modulator. The second readdressing block is a projecting block ("point-point"). The matrix

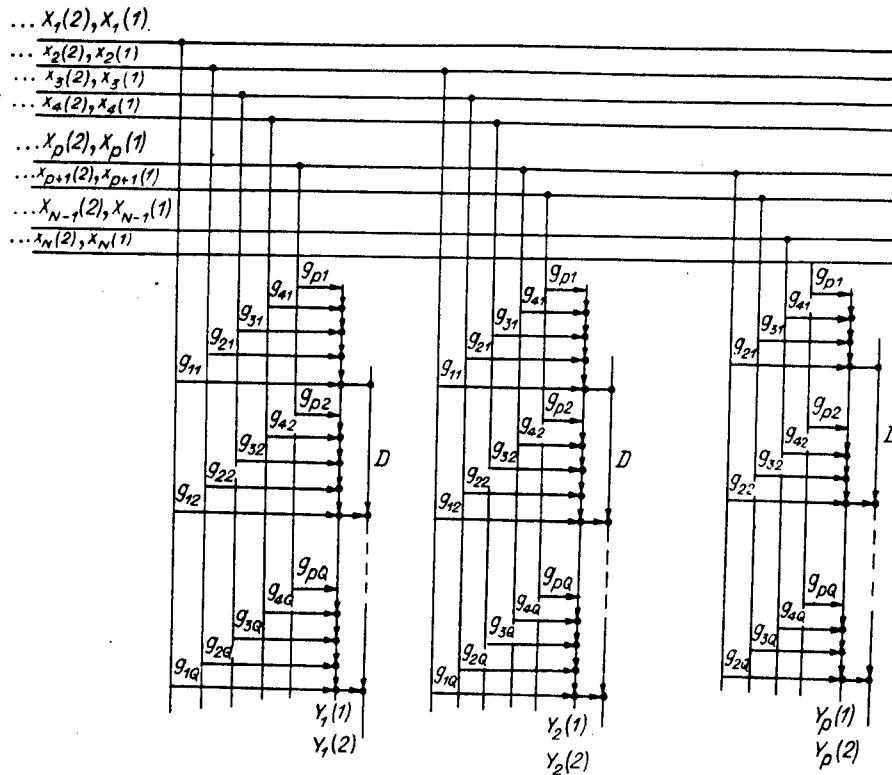


Figure 10.

multiplier in OSP philosophy is structurally one of the most elementary devices that has been suggested for this purpose.

§5. Tuning OSP for Two-Dimensional Convolution (Correlation) of Images [5]

The operation is described by

$$Y(k,l) = \sum_m \sum_n x(k-m, l-n) g(m,n).$$

The GPS of this transformation is illustrated by fig. 10. Its main features are the following: input signals are the lines of one of the images being formed; the input data indices are decremented by line scanning of the image; there is no decrement of coefficients index; only indices of output data are decremented; the input signal is a two-dimensional correlation function fed out line-by-line.

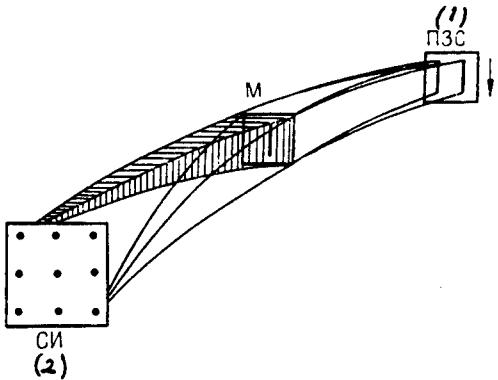


Figure 11.
Key: 1 — PZS; 2 — SL.

The tuning of OSP for two-dimensional correlation is illustrated by fig. 11. The light emitter format is a matrix. The first readdressing block transforms the flow from each light emitter onto the entire modulator region, with "point-plane" addressing. A mask is recorded on the modulator, whose transmittance is proportional to the kernel $g(m,n)$. The second readdressing block projects the modulator plane onto the photosensitive surface of the PZS-array, so that the image registered on the modulator and restored by the light emitter number m is shifted relative to the image restored by the first emitter to the distance $(m - 1)d$, where d is the image resolution element size. The type of addressing is "point-shifted points," where the shift is defined by the serial number of the light emitter that formed the point. With this addressing, a one-dimensional correlation function of the image line and the kernel registered on the modulator are computed. The correlation in the other coordinate is accomplished by the method usual for OSP: by decrementing output data indices (see §1).

In this paper we have demonstrated that the optoelectronic signal processor structure proposed makes it possible to reprogram a given basic structure for execution of a broad range of tasks in image and signal processing. In particular, we gave examples of tuning OSP for execution of five different functions.

Technological solutions executing possible adjustment of OSP for three of these five functions and experimental tests of these variants have been described in [6].

The accuracy of computations is known to be a major problem in analog computer technology. We will show that in OSP this problem is less acute than in other devices of this kind. The input signal in OSP is fully discretized in space, because it is formed by an ensemble of discrete

emitters. Likewise, the output signal is discretized, because it is read by a discrete device: the PZS-array. Usually, readdressing and lift flow modulation devices are also discretized. Thus, OSP is a computing device that is fully discretized in terms of space, i.e., a "mixing" of channels that deteriorates the parameters of analog computers is in this case ruled out. Discretization makes it possible also to drastically improve OSP accuracy by introducing bit-analog coding, i.e., representing output information in a binary form, processing each bit separately and obtaining the end-result by weighted summation of the results of processing of individual bits.

The future of OSP hangs on development of a component base. With the creation of line ranges and arrays of light emitters in an integrated form, OSP size could be greatly reduced. A major OSP unit is the addressing block. An optimal static addressing block would be based on optical fibers. In the future, a dynamic addressing block could be developed, for example, based on real-time holograms or a set of light-modulating keys. With such units it would be possible to reprogram OSP rapidly, thus expanding its functions. Progress in three-dimensional modulators in the near future would help rapid input of coefficient matrices or transformation kernels into OSP. Finally, a PZS-array which is a core element of OSP should also be designed to fit with the OSP structure. The operation speeds of OSP are currently limited mainly by data output speed from the PZS-array (~10 MHz). In a specialized array, this value could be improved by one or two orders of magnitude [7], with a proportional improvement of OSP speed. In the future, OSP could be conceived as a "sandwich" comprising a set of integrated elements: light emitters, modulators, PZS-structures and switches. These processors would be analogous to electronic VLSI TsSP [central specialized processors], but with an operation speed higher by two or three orders of magnitude than in electronic devices.

Bibliography

1. Nezhevenko, Ye.S., "Optoelectronic Signal Processors," in: TEZISY DOKLADOV I VSESOYUZNOY KONFERENTSII PO OPTICHESKOY OBRABOTKE INFORMATSII [Abstracts of Papers Presented at the First All-Union Conference on Optical Information Processing], Leningrad, LIAP, 1988.
2. Koul, B.K., "Digital Methods and Facilities in Signal Processing," Elektronika, No 18, 1985.
3. Sunhuan, G., "Systolic and Wave Matrix Processors for High-Performance Computations" (Russian translation) IEEE TRANSACTIONS, Vol 72, No 7, 1984.

4. Dagon, D., and Mercero, R., DIGITAL PROCESSING OF MULTIDIMENSIONAL SIGNALS, Moscow, Mir, 1988 (Russian translation).
5. Gofman, M.A., Korzhov, Ye.I., Nezhevenko, Ye.S., and Polezhayev, V.V., "Optoelectronic Pipeline Image Correlator," in: TEZISY DOKLADOV NA 6 VSESOYUZNOY SHKOLE-SEMINARE PO OPTICHESKOY OBRABOTKE INFORMATSII [Abstracts of Papers Presented at the Sixth All-Union Seminar-School on Optical Information Processing], Frunze, FPI, 1986.
6. Gibina, L.A., Gofman, M.A., Nezhevenko, Ye.S., and Feldbush, V.I., "Experimental Study of Optoelectronic Signal Processors in Image and Signal Processing Modes," AVTOMETRIYA, No 3, 1989.
7. TEKHNOLOGIYA PZS-INTEGRALNYKH MIKROSKHEM: OBZOR VINITI [PZS-Integrated Microcircuit Technology: A VINITI Survey], Moscow, 1986.

COPYRIGHT: "Nauka," "Avtometriya," 1989

UDC 681.382

Microelectronic Optical Digital Computer Devices

907G0029C Novosibirsk AVTOMETRIYA in Russian No 3 May-June 1989 (manuscript received 26 Dec 88) pp 61-68

[Article by V.M. Yegorov and E.G. Kostsov (Novosibirsk): "Microelectronic Optical Digital Computer Devices," published under the rubric "Organization and Structure of Optoelectronic Computer Systems"]

[Text] An essential improvement of operation speed of computer devices, raising it to a level of 10^{13} - 10^{15} and more unit operations per second, is the basic challenge facing modern microelectronics.

Historically, computer engineering has depended for a drastic performance improvement on the introduction of new element bases. An analysis of the current state of the element base in microelectronics suggests that connections on the surface of backings and between backings are the main obstacle to further progress.

In addition, the main reserve of performance improvement — increasing the clock speed of logic circuits — has been largely exhausted, because the switching time of elements has become comparable to the signal transmission time between blocks of the device, including the control block. The signal speed in integrated circuits is substantially slower than the speed of light (by a factor of 5-50, depending on the degree of integration).

We can say that the clock speed of computer devices is no longer determined by considerations of physical engineering but by circuit engineering limitations. In this context, the way to further improvement of computer performance appears to lie through enhancing the degree of parallelization of procedures at all levels, down to elementary operations. This should be accompanied by organization of the structure of connections and simultaneous increase in the number of logic elements. Yet, modern microcircuit technology is incapable of resolving the even more complex problem of connections that arises in this case.

One possible solution is by introducing optical communication channels into the structure of logical VLSI. We should note that development of optical digital devices based on new physical principles that would disregard the accomplishments of microelectronics would involve immense capital investment, comparable to what has been spent in developing the existing technological base of microelectronics. In addition, due to diffraction in propagation of light beams that imposes rigid constraints on the ratio of the communication channel length h and the light source size b (modulator window) — $b > \sqrt{h\lambda}$, where λ is the light wave length — logic elements have to be implemented as microelectronic units, otherwise the device would be unacceptably large. For instance, if the data exchange is conducted through free light rays between N arbitrary homogeneous elements situated on two different backings of size L , the relations $b > \sqrt{N\lambda}$, $L > N\lambda$ should be satisfied; a device with $N = 10^7$ elements would then take up a volume of >100 m³. These physical limitations make it necessary to largely rely in optical logic structures on algorithms operating according to the principle of information processing in a limited neighborhood of each element. Effective parallel data processing algorithms have already been created which are based on neighborhood principles.

Algorithms of generalized substitutions are a case in point [1].

Earlier we developed a philosophy of logical circuit construction where signals are transmitted through optical communication channels [2, 3]. The basic logic element that satisfies a set of requirements for logic elements consists of an electrooptical light modulator (the element's output), a photoelectric converter (the element's input) and an energy accumulator. With this element base we developed a concept of a functional element of a matrix processor [3], showing that, by introducing optical channels in logical circuit designs, one can reduce (by two orders and more) the total number of components in the device as compared with a purely electronic device that has an equivalent functional power. At the same time, the area occupied by connections can be reduced by almost an order of magnitude, while the performance can be increased (as applied to image processing) by a factor of m , where m is the dimension of the data file.

In the past few years this philosophy of computer design has been supported by other investigators [4, 5].

The objective of the present paper is to examine the specific features in the design of integrated computer devices based on the principles of electrooptical modulation of the light flow and to describe the optical switch element, which is a major unit in the computational structure: a communication channel switch.

At the present time, three physical effects should be considered for possible application in integrated technology of high-speed thin-film light modulators: the Franz-Keldysh effect [6], the effect of field shift of exciton absorption lines, including in multilayer superlattice structures [7,

8] and an electrooptical effect in ferroelectric materials [9]. It is a specific feature of the first two modulators that a light source with a narrow spectral emission range has to be used, and that they require a considerable increase of the conductivity of the electrooptical layer when illuminated (due to strong light absorption). The latter fact results in a complication of logic element structure and increases the number of its components.

Light modulators based on the Pöckels effect eliminate rigid constraints on the choice of light flow wavelength; the light flow energy is absorbed outside the electrooptical layer in the analyzer. For this modulator an additional analyzer layer has to be fabricated. Electronic lithography is capable of making an analyzer (a metal film on photodetector surface) as a polarizer grating with a step $d \approx 0.3\text{-}0.35 \mu\text{m}$ and transparency in the wavelength range $\lambda = 0.6\text{-}0.7 \mu\text{m}$. At $\lambda > 2d$, polarizer gratings provide a light wave polarization degree close to 100 percent and a transmittance of 60-90 percent [10]; the metal film thickness is $0.1\text{-}0.15 \mu\text{m}$.

The high speed of these light modulators can maintain a clock speed w of logic circuits in the range of 10-100 MHz and more. A limiting factor for increasing w is heat removal: it cannot exceed 10^5 W/m^2 (when the element is heated above 300 K by 20-30 K). An analysis of the operation specifics of a logic element suggests that, of two energy flows reading the logic element (the light flow energy and the energy of modulator capacity recharging C), the latter is greater at $w \gtrsim 10 \text{ MHz}$. Therefore, the main reserve for increasing w is in reducing the element supply voltage V_0 . One way of reducing V_0 is by using at the input of the logic element threshold photodetectors with appropriate choice of the working point of light modulator operation mode.

In order to demonstrate this possibility, we created on the basis of electrooptical crystals with half-wave voltage 180 V a model of a dynamic memory (the block diagram is given in fig. 1). It consists of two elements, marked by even and odd indices, respectively. Light modulators M_1 and M_2 form optical read signals I_{sch} ; the quarterwave elements $\lambda/4$ and analyzers A isolate the phase of optical signals arriving at photodetectors F , which are inputs of key circuits K .

During odd cycles, capacities C_1 and C_2 are charged alternately (information erasure); during even cycles information is read from outputs of modulators M_4 and M_3 into the inputs F_1 and F_2 of respective elements, thus accomplishing storage of the information unit in the cell. The information can be expressed as light signals or as voltages at light modulator electrodes (capacities C_1 and C_2).

It has been established experimentally that a steady operation of the dynamic memory is possible at $V_0 = 5 \text{ V}$; a further decrease in V_0 is prevented by

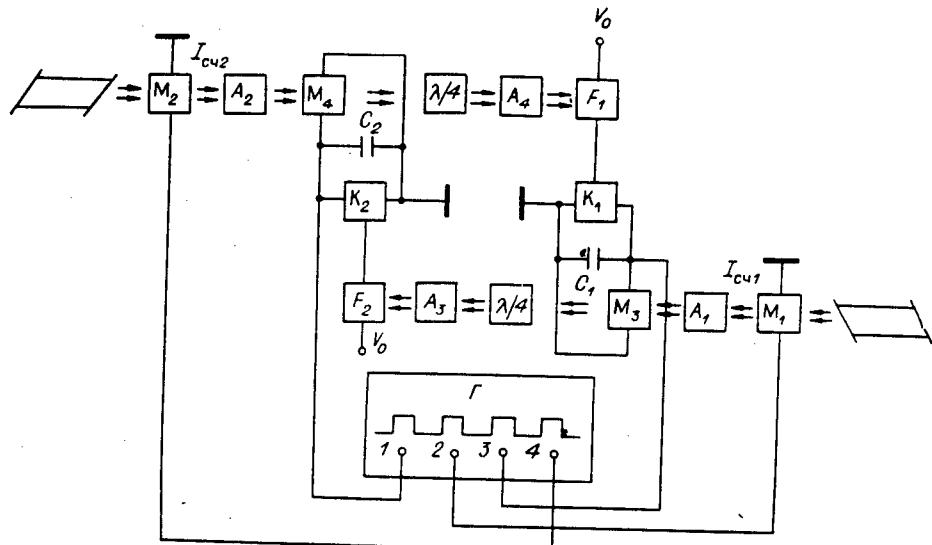


Figure 1.

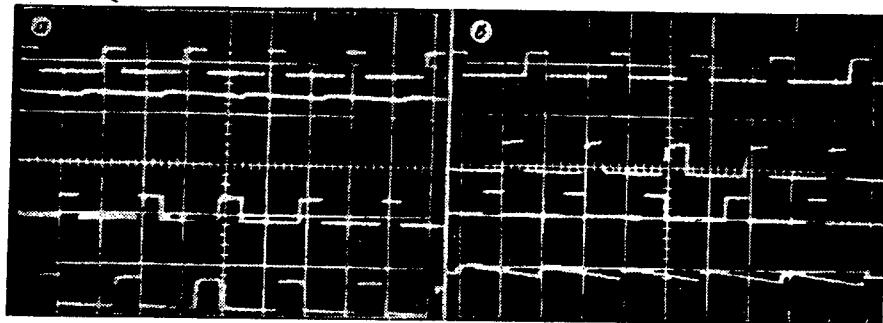


Figure 2.

noise from an He-Ne-laser used as the light source. Figure 2 shows typical oscilloscopes characterizing the time distribution voltages in the components of the device. The first and third beams describe the time behavior of the voltage in "even" and "odd" transparencies (accordingly, the behavior of intensity of the "reading" light flow is similar). The second and fourth beams indicate the change of the potential at the electrodes of light modulators of the "even" and "odd" logic elements.

Figure 2a illustrates the distribution of voltages where the registration of a logical unit was first performed in an "even" memory element; in fig. 2b it was first recorded in an "odd" element. The elements executed an inversion of the signal. The distribution of potentials set initially is retained for an indefinite time: it corresponds to an unattenuating propagation of

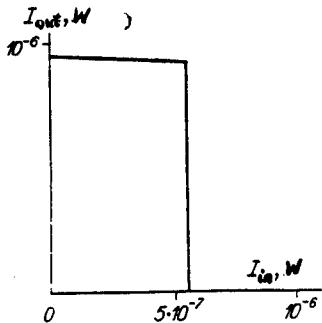


Figure 3.

information through an infinite string of elements. Signal quantization by each logic element maintained preservation of information even with certain variations of the photodetector illumination level, the cycle length τ and the amplitude of the control voltage V_0 .

We will estimate possible parameters of a logic element, proceeding from parameter values of its components that have been achieved as of the present time. The photodetector is a structure based on heterojunctions with steepness $S = 100\text{-}500 \text{ A/W}$, a reception area $30 \times 30 \mu\text{m}$, sensitivity 10^{-5} to 10^{-3} W and working frequency band up to 200 MHz [11]. The thin-film light modulator is based on one of the best electrooptical materials — barium strontium niobate with film thickness of $5 \mu\text{m}$ and dielectric constant $\epsilon = 1000$ [12, 13]; it has a capacity of $C \sim 4 \text{ pF}$ with a window area of $\sim 20 \times 20 \mu\text{m}$. Taking the limiting values of surface heat removal at 10^5 W/m^2 , we can easily estimate the cycle length $\tau > 10^{-7} \text{ s}$ and the switch energy 10^{-10} J at $V_0 = 5 \text{ V}$. Obviously, this energy is considerably larger than the light flow energy necessary to satisfy the photodetector sensitivity. The main reserve for reducing the energy intensiveness of these elements is thus in decreasing V_0 and ϵ . One way to reduce V_0 is by using superlattice modulators, where it is possible to reduce the capacity of an individual modulator down to $\sim 0.1 \text{ pF}$ and less, V_0 to $1.5\text{-}2 \text{ V}$ (taking into account the nonlinearity of photodetector characteristics) and the switching energy W to 10^{-12} J . Figure 3 shows the transfer characteristic of a logic element operating as an inverter constructed on the basis of a superlattice modulator and a threshold photodetector with a switching energy of 10^{-12} J at $\tau = 10^{-8} \text{ s}$, $C = 6 \times 10^{-14} \text{ F}$, $S = 500 \text{ A/W}$, $V_0 = 5 \text{ V}$; light flow intensity at the element input is $\sim 10^{-6} \text{ W}$ and light background is $\sim 2 \times 10^{-7} \text{ W}$. An element with this characteristic can support unattenuating propagation of information through an infinite string of elements.

Estimates of the limiting parameters of these logic elements indicate that the signal quantization property can be preserved up to $W = 10^{-13} \text{ J}$ and $\tau \approx 10^{-9} \text{ s}$. A further decrease in clock cycle length to 10^{-10} s and less in the design of large computational circuits will face serious circuit engineering difficulties which in our opinion are unresolvable.

In this context, we may note that in addressing this problem the prospects of using optical bistable elements based on nonlinear properties of certain semiconductor materials [14], recently put into the limelight because of their short switching times ($10^{-10}\text{-}10^{-12} \text{ s}$), are still questionable not only because of the above-mentioned difficulties, but also because there is no practical proof that it is feasible to create on their basis a functionally

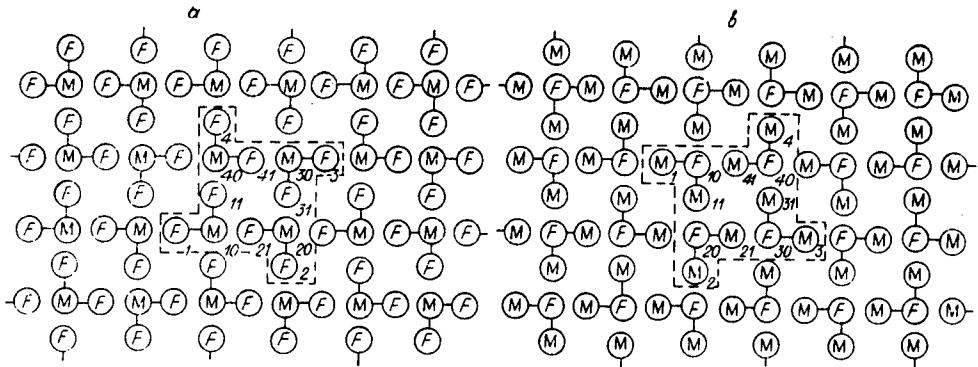


Figure 4.

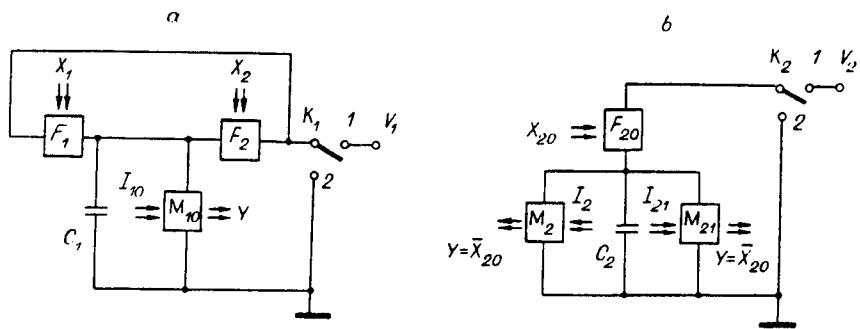


Figure 5.

complete logic element that would combine the properties of signal quantization and memorization, input merger and output branching, and technological efficiency in mass production.

We will consider a switching element with four optical inputs and outputs [15] constructed with the above element base and designed for switching four optical information channels in digital form.

An optical switching element [OKE] is comprised of 24 components: 12 light modulators and 12 photodetectors, forming two cells in two planes (fig. 4a,b). The first cell comprises four NOR circuits; the second cell includes four inverters with output splitting (fig. 5a,b).

We will describe the operation principle of the cell in a dynamic mode, as illustrated by a NOR circuit: cycle 1 — charging capacitor C , key K_1 is switched to the supply source V_1 ; cycle 2 — recording of information by optical signals arriving at photodetectors F_1 and F_2 , key K_1 in position 2; cycle 3 — reading information by optical signal from modulator M_{10} , key K_1 in position 2. If signals X_1 and X_2 arrive at the input of the photodetectors F_1

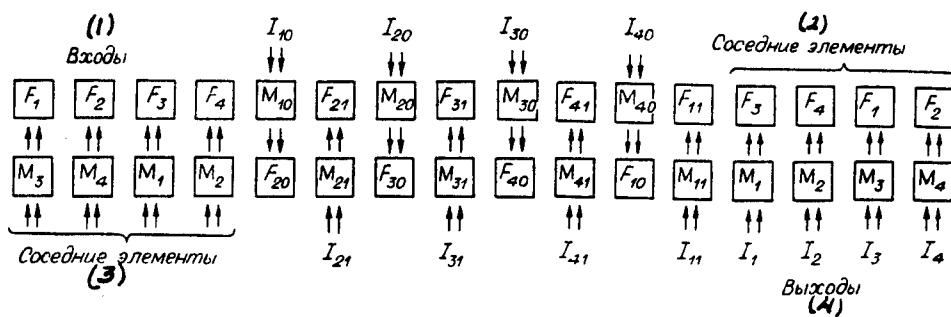


Figure 6.

Key: 1 — inputs; 2 — neighboring elements; 3 — neighboring elements; 4 — outputs.

Table 1.

Cycle number	f_{10}	f_{20}	f_{30}	f_{40}	f_{11}	f_{21}	f_{31}	f_{41}
4	1	1	1	1	0	0	0	0
6	0	0	0	0	1	1	1	1
8	1	1	1	1	0	0	0	0
10	0	0	0	0	1	1	1	1
12	1	1	1	1	0	0	0	0
14	0	0	0	0	1	1	1	1
16	1	1	1	1	0	0	0	0
18	0	0	0	0	1	1	1	1

and F_2 , the function $Y = (X_1 \vee X_2)$ is realized at the output of modulator M_{10} . The second cell operates similarly.

The reciprocal orientation of the cells is such that the output signals of modulators with double indices of a given cell are input signals for the respective photodetectors of the other cell; the output signals of modulators M_1 , M_2 , M_3 and M_4 are input signals for photodetectors F_3 , F_4 , F_1 and F_2 of adjacent OKE, respectively, and vice versa (fig. 6). Thus, OKE executes the functions of rotation and storage of four information units, transmission of these units to neighboring elements, or both simultaneously, depending on the combination of reading signals in OKE.

Table 1 specifies the function of storage and complete rotation, where f designates the binary function of the corresponding light flow 1. This assumes that information in OKE is recorded in the second cycle, while in

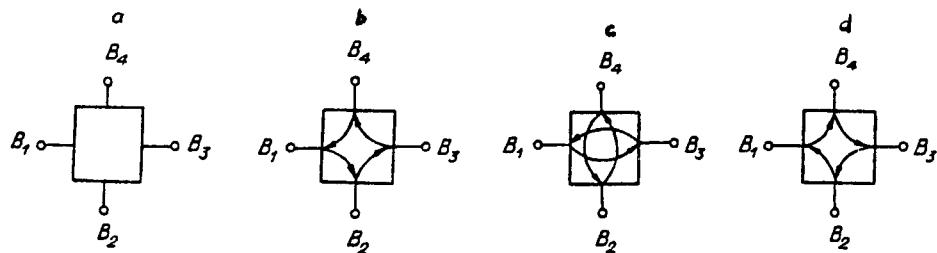


Figure 7.

each other cycle the capacity of the respective cell is charged. Therefore, the rotation vector $f_n = \{f_{10} f_{20} f_{30} f_{30} f_{40} f_{11} f_{21} f_{31} f_{41}\}$ has only two values: $f_{n_1} = (11110000)$ and $f_{n_2} = (00001111)$; the sequence of rotation to 90° consists of four cycles; the sequence of complete rotation consists of 16 cycles.

Table 2 presents two switching functions, where B_1, \dots, B_4 (fig. 7a-d) correspond to information channels with input and output $F_1, M_1, \dots, F_4, M_4$ of OKE; the channel switchings are indicated by arrows.

We see from table 2 that the switching type is determined by the instant of arrival at light modulators M_1, M_2, M_3 and M_4 of reading light flows I_1-I_4 : in each $(6 + 16n)$ -th cycle ($n = 1, 2, 3, \dots$) a switching of the type illustrated in fig. 7b is executed; in the $(10 + 16n)$ -th cycle switching of the type corresponding to fig. 7c; in the $(14 + 16n)$ -th cycle, fig. 7d.

The switching control vector $f_k = \{f_1, f_2, f_3, f_4\}$ in each case has 16 states; therefore, 46 methods of switching four communication channels are implemented. The splitting functions are formed by a combination of the three main switching types described in table 2.

In this optical switching element functional completeness is achieved not through rigid electric connections but by means of flexible optical connections; the absence in the design of this device of gates and switching elements, typical for purely electronic circuits, makes it much more compact than a similar device using the existing element base.

An optical switch can be created on the surface of a single backing (for instance, sapphire with a thickness of $50-100 \mu\text{m}$); with KNS [silicon on sapphire] technology, threshold photodetectors, electrooptical layers and electrodes (including those transparent to light) have been formed on such a backing. Proceeding from the energy demands for switching a single logic element estimated in the paper, the performance of this structure can be evaluated as $10^{14} \text{ Hz gates/cm}^3$

Table 2.

Cycle number	f_1	f_2	f_3	f_4	Switching type
$6 \pm 16n$	0	0	0	0	0
	0	0	0	1	$B_3 \dashrightarrow B_4$
	0	0	1	0	$B_2 \dashrightarrow B_3$

	0	1	0	0	$B_1 \dashrightarrow B_2$

	1	0	0	0	$B_4 \dashrightarrow B_1$

	1	1	1	1	$B_4 \dashrightarrow B_1, B_1 \dashrightarrow B_2, B_2 \dashrightarrow B_3, B_3 \dashrightarrow B_4$
$10 \pm 16n$	0	0	0	0	0
	0	0	0	1	$B_2 \dashrightarrow B_4$
	0	0	1	0	$B_1 \dashrightarrow B_3$

	0	1	0	0	$B_4 \dashrightarrow B_2$

	1	0	0	0	$B_3 \dashrightarrow B_1$

	1	1	1	1	$B_3 \dashrightarrow B_1, B_4 \dashrightarrow B_2, B_1 \dashrightarrow B_3, B_2 \dashrightarrow B_4$
$14 \pm 16n$	0	0	0	0	0
	0	0	0	1	$B_1 \dashrightarrow B_4$
	0	0	1	0	$B_4 \dashrightarrow B_3$

	0	1	0	0	$B_3 \dashrightarrow B_2$

	1	0	0	0	$B_2 \dashrightarrow B_1$

	1	1	1	1	$B_3 \dashrightarrow B_1, B_3 \dashrightarrow B_2, B_4 \dashrightarrow B_3, B_1 \dashrightarrow B_4$

Bibliography

1. Anishev, P.A., Achasova, S.M., Bandman, O.L., et al., METODY PARALLELNOGO MIKROPROGRAMMIROVANIYA [Methods of Parallel Microprogramming], Novosibirsk, Nauka, 1981.
2. Kostsov, E.G., and Mishin, A.I., "Specifics of Design of Optical Computers," MIKROELEKTRONIKA, Vol 6, Issue 2, 1977.
3. Yegorov, V.M., and Kostsov, E.G., "Prospects for Development of Optical Digital High-Performance Computer Devices," AVTOMETRIYA, No 1, 1985.
4. Lee, S.H., Esener, S.C., Title, M.A., and Drabik, T.J., "Two-Dimensional Silicon/PLZT Spatial Light Modulators: Design Considerations and Technology," OPTICAL ENGINEERING, Vol 25, No 2, p 250, 1986.
5. Esener, S.C., Wang, J.H., Drabik, T.J., et al., "One-Dimensional Silicon/PLZT Spatial Light Modulators," OPTICAL ENGINEERING, Vol 26, No 5, p 406, 1987.
6. Keldysh, L.V., "The Effect of a Strong Electric Field on Optical Characteristics of Nonconducting Crystals," ZHURNAL EKSPERIMENTALNOY I TEORETICHESKOYE FIZIKI, Vol 34, No 5, 1958.
7. Wood, T.H., Burrus, C.A., Miller, D.A.B., et al., "High-Speed Optical Modulation With GaAs/GaAlAs Quantum Well in a p-i-n Diode Structure," APPLIED PHYSICS LETTERS, Vol 44(1), p 16, 1984.
8. Miller, D.A.B., Chemla, D.S., Damen, T.C., et al., "The Quantum Well Self-Electrooptic Effect Device: Optoelectronic and Oscillation and Self-Linearized Modulation," IEE JOURNAL OF QUANTUM ELECTRONICS, Vol QE-21, p 1462, 1985.
9. Mustel, Ye.R., and Parygin, V.N., METODY MODULYATSII I SKANIROVANIYA SVETA [Methods of Light Modulation and Scanning], Moscow, Nauka, 1970.
10. Vanyurikhin, A.I., and Gerchanovskaya, V.P., OPTIKO-ELEKTRONNYE POLYARIZATSIONNYE USTROYSTVA [Optoelectronic Polarization Devices], Kiev, Tekhnika, 1984.
11. Zhuk, B.V., Elenko, A.A., Prokhorov, A.M., et al., "High-Speed Heterostructure Phototransistor," FIZIKA I TEKHNIKA POLUPROVODNIKOV, Vol 22, No 8, 1988.

12. Antsygin, V.D., Kostsov, E.G., and Sterelyukhina, L.N., "Impulse Electrooptical Light Modulation in Thin Ferroelectric Films," AVTOMETRIYA, No 5, 1983.
13. Antsigin, V.D., Egorov, V.M., Kostsov, E.G., et al., "Ferroelectric Properties of Thin Strontium Barium Niobate Films," FERROELECTRICS, Vol 63, p 235, 1985.
14. Smith, P.W., "On the Physical Limits of Digital Optical Switching and Logic Elements," THE BELL SYSTEM TECHNICAL JOURNAL, Vol 61, No 8, p 1975, 1982.
15. Kalyayev, A.V., ODNORODNYE KOMMUTATIONNYE REGISTROVYE STRUKTURY [Homogeneous Switching Register Structures], Moscow, Sovetskoye Radio, 1978.

COPYRIGHT: "Nauka," "Avtometriya," 1989

UDC 681.3.01:621.378

Architecture of an Information System Based on a Large-Capacity Holographic Memory

907G0029D Novosibirsk AVTOMETRIYA in Russian No 3 May-June 1989 (manuscript received 14 Dec 88) pp 74-82

[Article by B.V. Vanyushev, N.N. Vyukhina, I.S. Gibin, A.P. Litvintseva, T.N. Mantush, B.N. Pankov, Ye.F. Pen, A.N. Potapov, I.B. Tatarnikova and P.Ye. Tverdokhleb (Novosibirsk): "Architecture of an Information System Based on a Large-Capacity Holographic Memory," published under the rubric "Organization and Structure of Optoelectronic Computer Systems"]

[Text] Introduction

The major merits of an optical memory include the long time (tens of years) and reliability of information storage, the high recording density (up to 10^7 bit/cm 2) and the low storage cost (10^{-7} kopeck/bit). In addition, holographic memory is suitable for parallel information input-output.

With two-dimensional holograms a frame (a page) consists of 10^3 - 10^4 bits, which is a premise for fast associative searching and other forms of information processing in the optical and optoelectronic levels inside the memory (distributed processing). Accordingly, our holographic memory technology on nonreversible media is oriented toward applications in database computers and other information retrieval systems [1, 2].

At the moment, holographic memory applications in the practice of information systems are prevented by insufficient development of the element base (semiconductor lasers, controllable light transparencies, deflectors, matrix photoelectronic VSLI, etc.) and the absence of a perfect operational (reversible) carrier medium. However, for various applications (space research, geophysics, engineering, etc.) archival memories are needed which can be constructed on nonreversible photoregistration media. The crucial requirement for systems of this kind is ensuring long-term and reliable data storage with continuous data accumulation and slow change, such as in information reference systems and certain databases.

Systems with nonreversible media are usually constructed as modular systems that can handle data within one module (single-digit to double-digit megabytes) in real time. The "operational part" of data is either transferred to that module from other media or first buffered in the working environment, for example, on a magnetic disk for data collection and updating.

The Institute of Automatics and Electrometry of the Siberian Department of the USSR Academy of Sciences and the industry's scientific research institute have unique experience in development of holographic memory technology on flat carriers (modules made of photographic plates on a glass backing with dimensions $76 \times 76 \times 2.65$ mm³). An archive has been created which can accommodate 5-10Gb of digital data and hundreds of thousands of frames of document data in a nondigital original form. Existing holographic ZU [memory devices], the optical aspects of the technology and the programs created for studying them are described in [3-8].

Here we will describe the functional organization (architecture) of an information system based on holographic ZU — a complex of hardware and software executing control of devices and data — and report the results of studies and applications with this set of programs.¹

Figure 1 depicts the components and general structure of an experimental holographic memory system with an SM-4 computer complex. Holographic ZU [GZU] representing digital and documentary archives are connected to a UNIBUS highway. These are optomechanical devices for registration [OMUZ] of holograms with data in digital and documentary forms (hereafter, "digital and documentary holograms"), optomechanical devices for reading and searching data [OMUChP] of digital holograms and for reading [OMUCh] documentary data with an electronic system for control of writing, reading and searching.

The hardware complex for input, representation, registration and documentation includes standard devices of a computer complex (NML [magnetic tape memory], NMD [magnetic disk memory], alphanumeric display terminals [ATsD0-ATsD3] and ATsPU [alphanumeric printing devices]) and, in addition, black-and-white and color video monitors (controlled by DINAMO and VIDEObufer), which provide for display of graphics information and halftone images, devices for document microcopying and photostat copying.

Hardware

We will describe the hardware of the experimental digital system as one with a more complex organization and structure than the component for documentary information.

The main components of the optomechanical devices OMUZ and OMUChP are assemblies of two-coordinate positioning of modules and selection of a module

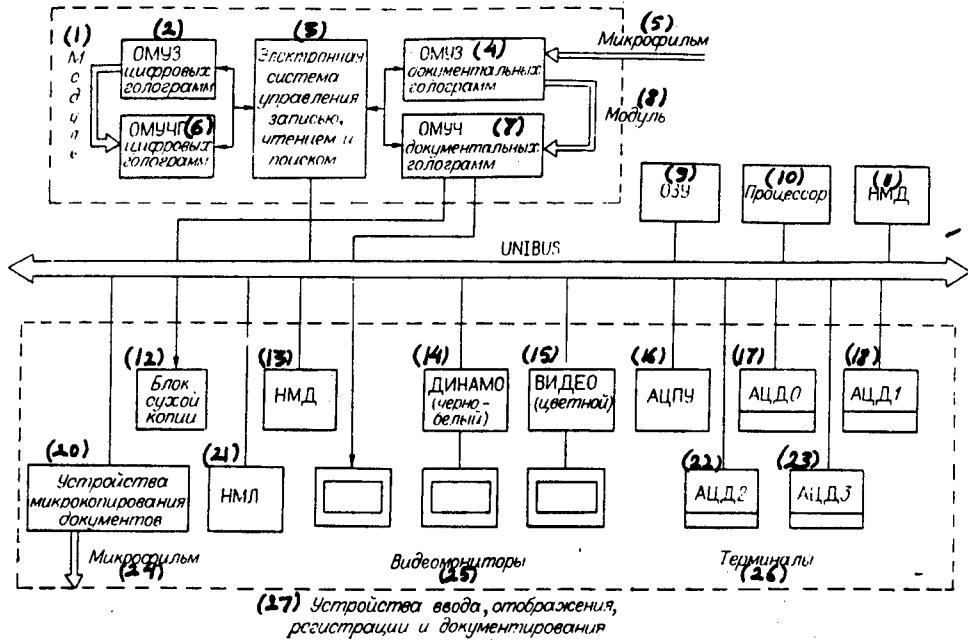


Figure 1.

Key: 1 — module; 2 — OMUZ of digital holograms; 3 — electronic reading/writing/searching control system; 4 — OMUZ of documentary holograms; 5 — microfilm; 6 — OMUChP of digital holograms; 7 — OMUCH of documentary holograms; 8 — module; 9 — OZU; 10 — processor; 11 — NMD; 12 — photostat block; 13 — NMD; 14 — DINAMO (black and white); 15 — VIDEO (color); 16 — ATsPU; 17 — ATsD0; 18 — ATsD1; 20 — document microcopying device; 21 — NML; 22 — ATsD2; 23 — ATsD3; 24 — microfilm; 25 — video monitors; 26 — terminals; 27 — input representation, registration and documentation devices.

from an archive (during reading), a laser, a controlled transparency forming 32-by-32 bit data pages, a camera shutter providing required exposure time, Fourier lenses, matrix LSI photodetector for conversion, processing and reading of reconstructed page images, deflectors for rapid hologram selection and for maintaining (correcting) the beam at the hologram center.

The electronic devices of the system control the writing and reading of information pages (holograms) with addressed access, associative search for data at the optoelectronic physical level and communication with the working computer. They constitute the three functional subsystems (otherwise referred to as controllers): WRITING, READING and SEARCHING.

The electronic subsystems (fig. 2) are built on the basis of microprocessors [MKP] from Elektronika-60 computers which control specialized writing and reading devices connected to the Q-BUS line. These subsystems include 64K working memories used by programmed processors of physical control of writing, reading and searching, and terminal interfaces [IT]. Terminals operate with control processors and with the operating system of the working computer.

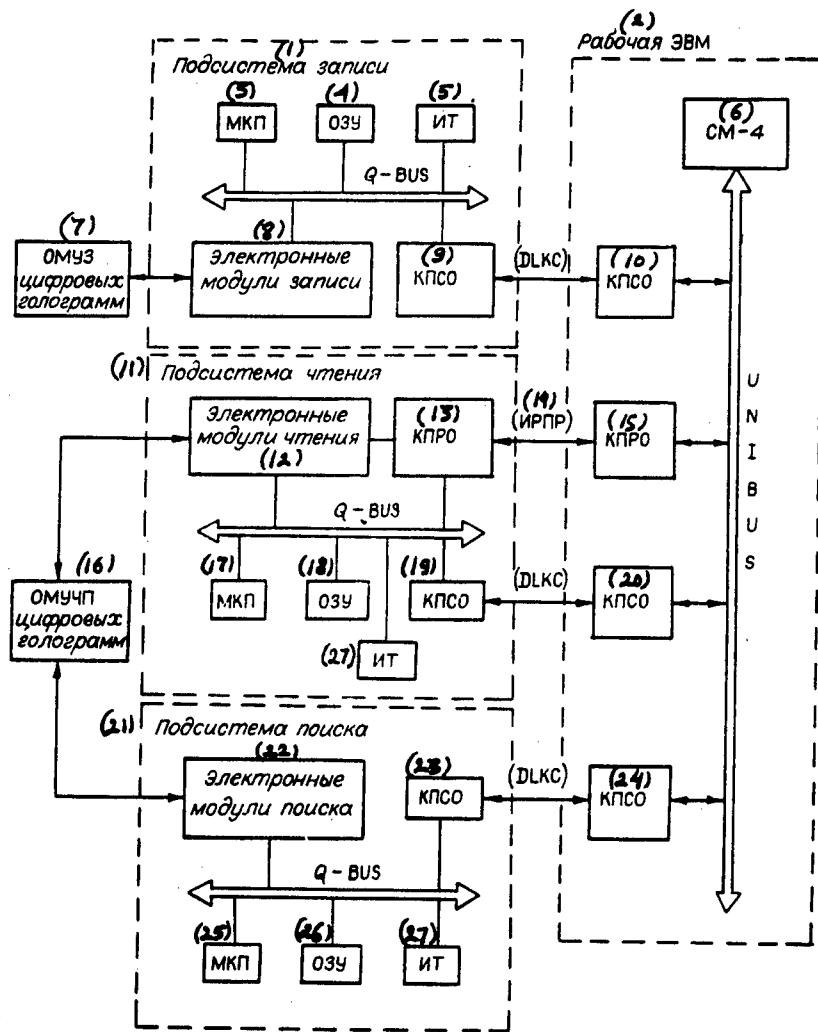


Figure 2.

Key: 1 — writing subsystem; 2 — working computer; 3 — MKP; 4 — OZU; 5 — IT; 6 — SM-4; 7 — OMUZ of digital holograms; 8 — electronic writing modules; 9 — KPSO; 10 — KPSO; 11 — reading subsystem; 12 — electronic reading modules; 13 — KPRO; 14 — IRPR; 15 — KPRO; 16 — OMUCHP of digital holograms; 17 — MKP; 18 — OZU; 19 — KPSO; 20 — KPSO; 21 — searching subsystem; 22 — electronic search modules; 23 — KPSO; 24 — KPSO; 25 — MKP; 26 — OZU; 27 — IT.

Sequential DLKC interfaces (sequential communication controllers [KPSO]) connect the working computer with the writing and reading subsystems, load programmed processors of physical writing/searching control from the disks of the working computer into the working memory and support the operation of remote terminals.

The READING subsystem utilizes both the sequential channel (DLKC) and a parallel channel (controllers of parallel communication [KPRO] with an IRPR interface). The former is used to transmit instructions to the controller

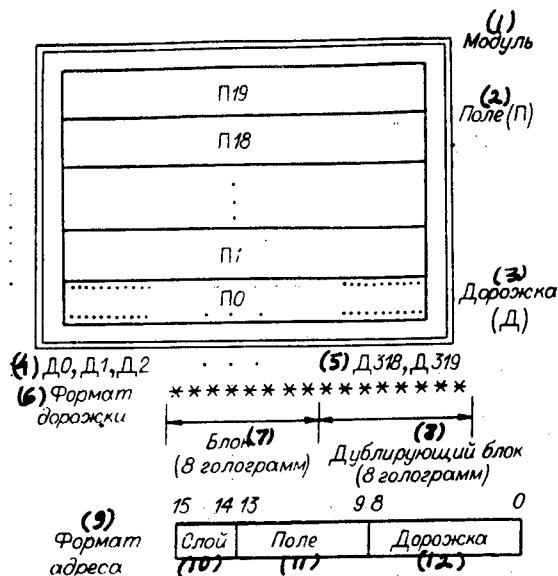


Figure 3.

Key: 1 — module; 2 — field (P); 3 — track (D); 4 — D0, D1, D2; 5 — D318, D319; 6 — track format; 7 — block (8 holograms); 8 — duplicate block (8 holograms); 9 — address format; 10 — layer; 11 — field; 12 — track.

field and a track is accomplished by moving the module; access to a hologram (a page) is accomplished by an acoustooptical light deflector.

A datablock that occupies half of a physical track consists of eight holograms. In order to enhance reading reliability, the record of a block is duplicated in the second half of a track (see fig. 3).

The data address format is this: layer → field → track.

For verification of address in reading the address in the above format is placed at the beginning of each block.

The block length is determined by the method of data coding on a page; for a Hamming code it is 672 bytes; for paraphase code, 512 bytes (including 2 bytes of block address).

The data on a page with a systemic Hamming code (22,16) are packed so that a word is accommodated in three sequential bytes: two information bytes and one checkbyte. The last two bytes contain a "unit marker" that can be used to automatically select the reading mode depending on the page image intensity. Data in the reading subsystem are represented in Hamming code.

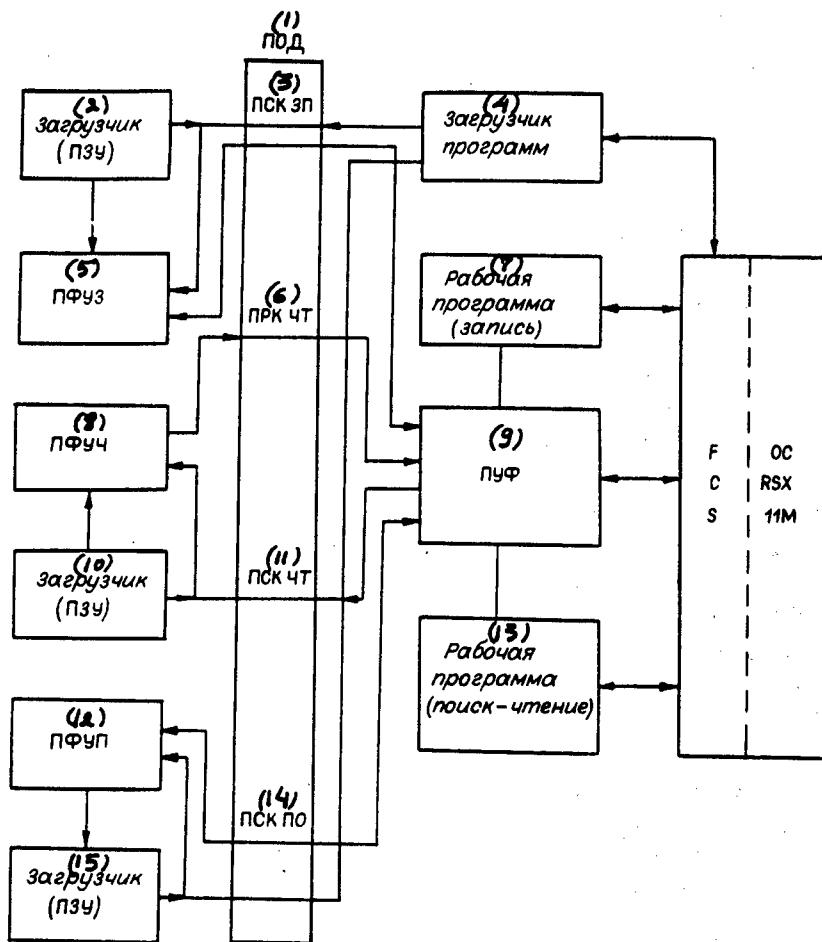


Figure 4.

Key: 1—POD; 2—loader (PZU); 3—PSKZP; 4—program loader; 5—PFUZ; 6—PRKChT; 7—working program (writing); 8—PFUCh; 9—PUF; 10—loader (PZU); 11—PSKChT; 12—PFUP; 13—working program (searching/reading); 14—PSKPO; 15—loader (PZU).

Data in the searching subsystem are written in paraphase code (0:10, 1:01), and a page is packed with two-byte words.

Software System

The main functional components of the program controlling devices and data (fig. 4) are the processors of physical control of writing [PFUZ], reading [PFUP] and searching [PFUP], file control processor [PUF] and the processor for exchange of data [POD].

The PFUZ, PFUCh and PFUP processors are implemented for the Elektronika-60 microcomputer family. They control optomechanical, optoelectronic and

electronic memory devices. They are stored on working computer disks and are loaded by a special PROGRAM LOADER (from the working computer) and hardware loaders of respective controllers.

The POD processor is formed by a distributed program which executes exchange of data between the working computer and the controllers through corresponding sequential channels (PSKZP, PSKChT and PSK PO) and a parallel channel (PRKChT). In the working computer the POD element processor employs terminal drivers of the operating system (the PROGRAM LOADER also works with them); in the controllers it uses special communication subprograms. The POD element on both sides verifies communication by the checksum of blocks, reports to the programs the status of the communication and sends an error message.

The PUF processor executes maintenance of user files in a module. In its operation it utilizes the capacities of a file control system (FCS) of the working operating system (RSX11M).

All the processors are built as a set of object modules of subprograms corresponding to commands. For executing a command subprograms must be called according to the JRS PC instruction, IMYa KOMANDY [command name]; return is accomplished according to the RTS PC instruction. The exchange of arguments is carried out through parameter blocks created in the user program and specified in the register R5.

We will briefly describe the commands of programmed processors.

The PFUZ processor executes two commands: INIM - module initiation for writing; and WRID - data block writing under address.

The INIM command has the following parameters: the number of the layer from which the writing is to be performed; the type of data coding on the page; the exposure energy for hologram registration; the step of writing specifying the multiplicity of the distance between neighboring holograms; and two-coordinate displacements of hologram matrices from edges of the module. The command positions the module to the initial point, determines the exposure time required for the specified energy and initiates the hardware-software elements for time corrections (in response to laser power drift).

The WRID command includes the logical number of the block that specifies the address of its record, the length of the datablock and the block itself. The PFUZ processor converts the block number into its physical address, positions the module, uses a transparency to form a two-dimensional page image from the datablock and executes the recording.

The PFUCH processor carries out four commands: INIR - initiate module reading; SELM - select the module from the archive; READ - read the block at the address; RETM - return the module to the archive.

The INIR command sets into the initial state all the elements of the unit, selecting and receiving the module from the archive and the reading mode: the accumulation energy for the photoarray, the threshold of its signal amplifiers that sets the boundary for separating 1 and 0 on the page, the reading step (on analogy with the writing step) and the number of the reading channel (the archive includes two channels).

The SELM command contains the channel number and the address of the module in the archive. The command causes the module to be selected from the archive and placed into the channel positioner. The photomatrix accumulation mode is determined, and the correction of the position of the reading laser beam (adjustment) with respect to the hologram center is accomplished.

The READ command, which contains the logical number of the block, reads the block and transfers it into the computer. In case of double Hamming error, the duplicate block is read, which may result in a change of the reading mode (the resetting of amplifier thresholds) for the duplicate and main blocks. Thus, the READ command can activate effective tools for correct reading of a data page. During the course of sequential reading, the reader beam is adjusted at transition to a new hologram field and after reading a certain set number (several dozen) of tracks in a given field.

Under the RETM command (with channel number and address as parameters) the module from the channel specified is returned to the archive under the given address.

The PFUP processor has five commands: INIC - searching channel initiation; REBL - datablock reading; RERC - record reading; FICO - search and counting of records according to a condition; FINR - search for numbers of records according to a condition.

The INIC command sets to the initial state the devices of the searching subsystem and the operation mode of the photodetector associative device.

The REBL command reads the datablock with the logical number specified.

The RERC command has these parameters: logical number of the file beginning block and the length, serial number and quantity of sequentially placed records. The beginning and size of the region of records to be read are determined from these data.

The FICO and FINR commands include parameters for RERC (which specify the search region) and the conditions for the search, which determine the

comparison operation (equal to, less than, greater than or a range) and a key (its code, size and shift in the record). The commands inspect a specified region of records according to the search condition; FICO returns the number of records that satisfy the search condition, and FINR returns the quantity and serial numbers of these records.

As the PFUP processor executes commands it checks for errors in optoelectronic data processing at the physical level. In case of an error, hologram selection and processing are repeated in the main field and, if necessary, in the backup field. The PFUP processor returns codes of command execution status.

The next (higher) level of the program is the file level. Description of the file system developed for the GZU is beyond the scope of the present paper. We will simply list the commands and their purposes. The commands of the module file control processor include: OMPW — open the layer of the module for writing; OPFW — open file for writing; WRIB — write a virtual block; PUTS — place a record in sequential access mode; CLFW — close file after writing; CLMW — close module after writing; OPM — open module layer (for reading or searching); OPF — open file (for reading or searching); REAB — read virtual file block; GETS — read record in sequential access mode; GETR — read record (of fixed length) in random access mode; FICN — count records satisfying a condition; FIRN — find serial number of records satisfying a condition; CLF — close file (after reading or searching); CLM — close module (after reading or searching).

Obviously, the set of commands usually employed in such processors has been supplemented by commands of associative data search which, as mentioned earlier, are maintained in the hardware.

Results of Studies

The set of studies that have been carried out included development, analysis and experimental testing of methods of hologram writing and reading capable of ensuring the required data reading reliability. The effects have been studied of the instability of the laser and the controlled transparency, the defects of the carrier medium and the influence of defects on the characteristics of holograms and reconstructed page images [3, 5, 7, 8].

Test programs were written expressly for this study which run the processors of physical control of the devices.

The hardware-software adjustment of exposure time in writing (laser drift) and accumulation time of optical signals by photomatrix in reading (scatter of intensity of the pages reconstructed), the spatial adjustment of the reading beam relative to the hologram center after reading several tens of

tracks, the duplication of each block in the track and the use of the Hamming code in the reading subsystem, which corrects single errors and detects double errors (on the length of 22,16) made it possible to reduce the probability of error in the reading of module down to a level of 10^{-9} .

Paraphase coding in the searching subsystem reduces the scattering of diffractional efficiency of holograms restored and makes it possible to employ hardware facilities for adjustment of accumulation time and error control.

The information system proved to be suitable for storage of real data and for experimentation with associative optoelectronic information searching. Digital data were entered into the system from magnetic media (NML and NMD).

One memory module with the use of the Hamming code can accommodate up to 5000 pages of text (3500 characters per page) or about 67 monochrome image frames of 512×512 pixels with 256 pixel gradations (for color images, 22 frames). With paraphase coding the information capacity of a module is 20-25 percent smaller.

The searching system was tested on a group of files that made up a small database for different-node fragments of chemical structures with a total volume of some 1Mb (a total of 13 files). All files consist of constant-length records containing information fields and fields indicating links between file records. The files of records with variable-length fields (chains of symbols, recurring groups and multivalued fields) were transformed to fixed-length record file ensembles with the use of link indicators.

In the searching subsystem a method of physical arrangement of records of tabular files was implemented (it will be referred to as the page-bit method); it is aimed at parallel associative searching with a special LSI with parallel optical input [9].

The records in a file are interpreted as the rows of a table; the record fields as table columns. A fragment of a table file is mapped directly into a two-dimensional physical page of the memory (a hologram) according to its size of 32×32 bits (32 rows of 2 bytes each in a paraphase code). A memory page thus can contain 2-byte fragments of like-named fields from 32 records that can be processed in parallel; each record in the general case is stored in like-named rows of several consecutive pages [10].

It should be clear that page-bit organization and the arrangement of records support a simple implementation of not only parallel comparison in any field (column) but also direct access to records. During a search a group of pages of 32 records (a segment) is interrogated for a match to a key submitted and a comparison statement; if a positive match is obtained, the serial numbers of the records that satisfy the request are calculated. Since the system

scans only those pages that contain the required record (selection) or record field (comparison), the access time and data volume processed are greatly reduced.

Conclusions

The system has been subjected to a set of successful physical experimental studies of devices of promising systems of optical (holographic) memory: methods and modes of data page writing and reading have been elaborated, the information errors that may occur in the process have been analyzed and methods have been found for eliminating them to achieve an acceptable reading reliability.

A basic software has been created for a holographic memory system which comprises processors of physical control of devices and a file system. Users are given a set of commands that they can employ in their own programs when working with a holographic memory.

At the file level information is stored in the form of texts, graphs and images. An experimental database model with parallel optoelectronic associative information search has been created which handles structural chemical formulas. Due to hardware-based searching, it is possible for an average file of 1000 records of tens of bytes each to reduce the number of page accesses by an order of magnitude compared with the conventional method of computer input and processing of records.

Footnote

1. Experiments were conducted and hardware was developed with the participation of A.A. Blok, V.A. Dombrovskiy, S.A. Dombrovskiy, A.V. Volkov, K.B. Tyunyukov, V.Ye. Butt, V.I. Kozik, V.D. Barmasov, S.M. Bechashnov and V.I. Shkuratov.

Bibliography

1. Moto-Oka, T., editor, FIFTH-GENERATION COMPUTERS: CONCEPTS, PROBLEMS AND PROSPECTS, Moscow, Finansy i Statistika, 1984 (Russian translation).
2. Danilov, G.A., and Shmid, A.V., "Database Machines," ITOGI NAUKI I TEKHNIKI. SERIYA TEKHNICHESKAYA KIBERNETIKA, Vol 22, 1987.

3. Blok, A.A., Vanyushev, B.V., Volkov, A.V., et al., "A Device for Automatic Registration of Matrices of Digital Data Holograms," AVTOMETRIYA, No 3, 1984.
4. Vanyushev, B.V., Volkov, A.V., Gibin, I.S., et al., "A Device for Storage and Reading of Digital Data in a Holographic Archival Memory System," AVTOMETRIYA, No 3, 1984.
5. Vyukhina, N.N., Litvintseva, A.P., Mantush, T.N., et al., "Programming Tools for Holographic Memory Analysis, Debugging and Testing," AVTOMETRIYA, No 3, 1984.
6. Blok, A.A., Vanyushev, B.V., and Volkov, A.V., "Archival Holographic Memory With Capacity of 1Gb," in: TEZISY DOKLADOV II VSESOYUZNOY KONFERENTSII PROBLEMY RAZVITIYA RADIOOPTIKI [Abstracts of Papers Presented at the Second All-Union Conference on Problems of Development of Radiooptics], Tbilisi, 1985.
7. Blok, A.A., Dombrovskiy, V.A., Dombrovskiy, S.A., and Pen, Ye.F., "Experimental Studies of Data Reading Reliability in Holographic Memories," AVTOMETRIYA, No 3, 1984.
8. Blok, A.A., Dombrovskiy, V.A., Dombrovskiy, S.A., et al., PRAKТИЧЕСКИЙ ПРЕДЕЛ ПЛОТНОСТИ ЗАПИСИ ДАННЫХ В ГОЛОГРАФИЧЕСКИХ ЗУНАХ ПЛОСКИХ НОСИТЕЛЕЙ [Practical Limit of Data Recording Density in Holographic Memories on Flat Media], Preprint of the Siberian Department, USSR Academy of Sciences, Novosibirsk, 1988.
9. Konyayev, S.I., Kashcheyev, E.L., and Pankov, B.N., "Photomatrix Associative Memory - Multifunctional Optoelectronic Element: Characteristics and Applications," in: TEZISY V VSESOYUZNOY KONFERENTSII PO GOLOGRAFIYI [Abstracts of Papers Presented at the Fifth All-Union Conference on Holography], Riga, 1985.
10. Butt, V.E., Vyukhina, N.N., Kozik, V.I., et al., "The Search for Compounds by the Fragments of Structural Formulae in Holographic Storage," in: THE THIRD USSR-JAPAN SYMPOSIUM ON COMPUTER CHEMISTRY (20-24 JULY 1987), Novosibirsk, 1987.

COPYRIGHT: "Nauka," "Avtometriya," 1989

UDC 681.327.68:621.373.826

High-Speed Digital Data Memory on an Optical Disk Pack

907G0029E Novosibirsk AVTOMETRIYA in Russian No 3 May-June 1989 (manuscript received 27 Dec 88) pp 82-94

[Article by Yu.V. Vovk, L.V. Vydrin, N.N. Vyukhina, V.N. Zatolokin, P.Ye. Tverdokhleb, I.Sh. Shteynberg and Yu.A. Shchepetkin (Novosibirsk): "High-Speed Digital Data Memory on an Optical Disk Pack," published under the rubric "Organization and Structure of Optoelectronic Computer Systems"]

[Text] Extra large digital data flows (10-100Gb) arriving at a rate faster than 100Mb/s (fiberoptic communication lines, digital television, high-performance computer complexes, etc.) present a difficult dilemma: on one hand, an information volume of 10Gb is greater than the maximum capacity of a single magnetic or optical disk; on the other hand, a single-channel optical registration device can operate at a maximum speed of 30Mb/s.

The conventional solution in magnetic registration is to use a pack of disks. The required memory capacity is then achieved by increasing the number of disks, while rapid registration is accomplished with several heads working in parallel.

In developing a similar memory with an optical disk pack, the basic problem is creating an optical head of a small size and weight that would provide a recording speed of at least 30 Mb/s with a data packing density of $\geq 2 \times 10^5$ bits/mm².

We have proposed for this purpose a digital data registration method utilizing one-dimensional linear Fourier holograms [1]. We have studied methods of frequency synthesis of such holograms using an acoustooptical modulator [AOM] and registration by emission of a semiconductor laser [2] and parallel heterodyne data reading [3] with multilevel relative phase information coding in the hologram.

The objectives of the present paper are: to develop methods of raising the speed and registration density that have been suggested and improve the specifications of the basic ZU [memory] element — the high-speed optical

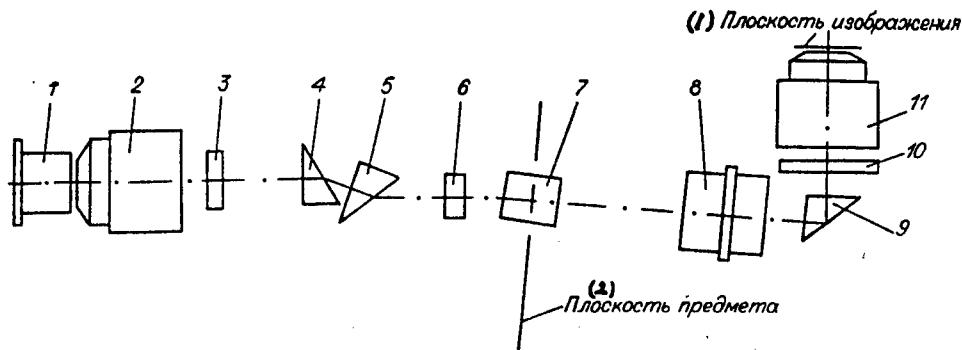


Figure 1.
Key: 1 — image plane; 2 — object plane.

head for binary information writing and reading; to work out the principles for organization and implementation of an experimental memory in a pack of optical disks with the following characteristics: capacity of 10Gb and registration speed of $\geq 128\text{Mb/s}$; to conduct an experimental study of the memory and the optical head in writing and reading modes.

Optical Head for Binary Information Writing-Reading

The head is depicted in fig. 1. It comprises an ILPN-108 semiconductor laser 1, a collimator-microlens $\times 40$ 2, a plate rotating the light polarization plane to 90° 3, prism collimators 4,5, a cylindrical lens 6, an acoustooptical modulator 7 [4], an Industar-M objective 8, a rotation prism 9, a block of two plane-parallel plates 10 and a microlens $\times 60$ 11.

The optical component of the reading channel (not shown in fig. 1) consists of light fiber bundle (lightguide) situated on the opposite side of the disk and an FPZ-4 photodiode installed on the reading amplifier board and connected optically to the head through the lightguide.

The head operates on the basis of the method of frequency synthesis of one-dimensional Fourier holograms by means of an acoustooptical modulator [2] and parallel heterodyne information reading [3].

A light beam from the semiconductor laser 1 is collimated by the microlens 2 to a parallel light beam (in the direction perpendicular to the p-n-junction plane of the laser). The beam is transmitted consecutively through plate 3, which rotates its polarization plane 90° , and the prism system 4,5, which reduces the light beam in the direction of the normal to the laser p-n-junction plane. The collimated beam is then focused by the cylindrical lens 6 to the zone of ultrasound wave of the modulator 7 (the object plane). The optical projection system, which consists of cofocally situated lenses 8

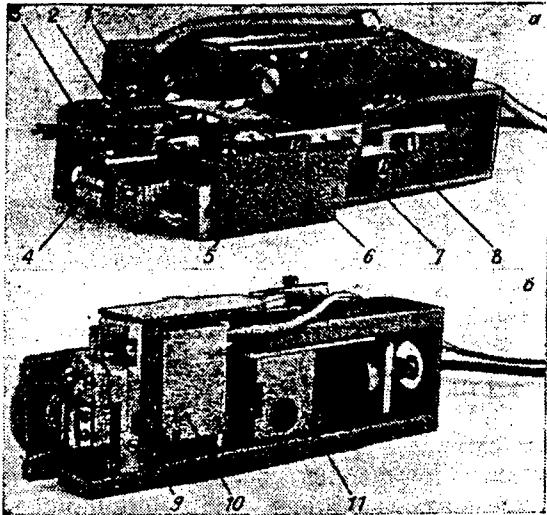


Figure 2.

and 11 and prism 9, carries light beams emerging from acoustooptical modulator 7 into the plane of the optical disk (the image plane) and forms in this plane a hologram which presents a reduced image of a visualized acoustic wave.

Due to anisotropic diffraction in AOM, the polarization plane of diffracted light waves is turned 90° relative to the incident wave. For obtaining a high contrastivity of the image of gratings, this mismatch must be compensated for. The directions of polarization planes are equalized by two phase rotating plates 10. The plates are installed in such a manner that all diffracted light beams are transmitted through one of them without changing the direction of the polarization vector; the zero beam is transmitted through the second plane with rotation of the polarization plane to 90°. As a result, the polarizations of reference and signal light waves in the area of the registration medium coincide. The design of the optical head in two positions is shown in fig. 2a,b. The head is comprised of functionally independent assemblies: head body 5, illuminator 11, acoustooptical modulator 10, the aerostatic suspension of microlens 2 and the reading block 1. Inside the head body these assemblies are mounted: modulator-illuminator, Industar-M objective 9 and rotating prism with the block plane-parallel plates 4.

The illuminator 11 comprises a semiconductor laser, the collimating microlens, the plate rotating the light polarization plane, a prism system and a cylindrical lens. The illuminator components are mounted in a separate housing. After they are completely adjusted, the acoustooptical modulator 10 is attached to this housing. The aerostatic suspension assembly consists of an air support plane 3, where the microlens 2 of projection optics is

mounted, the spring 6 and the device removing and locking suspensions 7,8 when the head is moved out of the working zone of the optical disk pack.

A Model of Optical Disk Pack Memory

Figure 3 gives the flowchart of a device for registration of binary information in the form of one-dimensional Fourier holograms.

From the voltage of the reference generator 1, N frequency-equidistant harmonics are formed in the unit of parallel synchronous frequency synthesizer 2; they are synchronized within the value of slowly varying phase φ_i , i.e., $U_i = U_m \cos[(\omega_0 + i\Delta\omega)t + \varphi_i]$.

In order to improve noise immunity, speed and registration density, a four-level relative phase information coding in the hologram is used. The information parameter in this data representation is the phase difference of the same space harmonic of two neighboring holograms [1, 5]. With four phase gradations it is possible to write a two-bit binary word on each space frequency.

The modulator 3 modifies the phase of each harmonic, so that its increment corresponds to the group of bits being recorded [5]. The voltages of all N harmonics are added up in linear summer 4 and are then fed to the electrical input of AOM 5. The parameters of AOM 5, the laser 6, the collimator 7 and the cylindrical lens 8 are such that, after a time $\Delta t = n2\pi/\Delta\omega$ ($n = 2$), the signal is changed completely in the modulator's light aperture. A brief current pulse is then sent to the laser from the supply power unit 9, and by means of a telescopic system 10,11 the image of the hologram (a set of sinusoidal gratings) is transferred to the light-sensitive coating of a disk 12.

When information is read from a hologram, the semiconductor laser is switched to continuous emission mode. A voltage is fed to AOM which is equal to the sum of all the harmonics, and the image of N moving sinusoidal gratings is projected onto hologram 13. As they interact with space harmonics of the hologram, a multifrequency signal appears at the outputs of photodetector 14 and amplifier 15. For each of the harmonics of this signal the phase shift between the voltage fed to the AOM and the voltage at the photodetector output depends on the phase shift of the respective hologram grating relative to the moving reading grating.

Obviously, the phase value is susceptible random fluctuations because of wobbling of the mechanical component of the memory, temperature variations and a drift of parameters of the optoelectronic channel. However, we observed in experiments that the typical fluctuation length along a track is a few millimeters and characteristic fluctuation time is hundredths of a

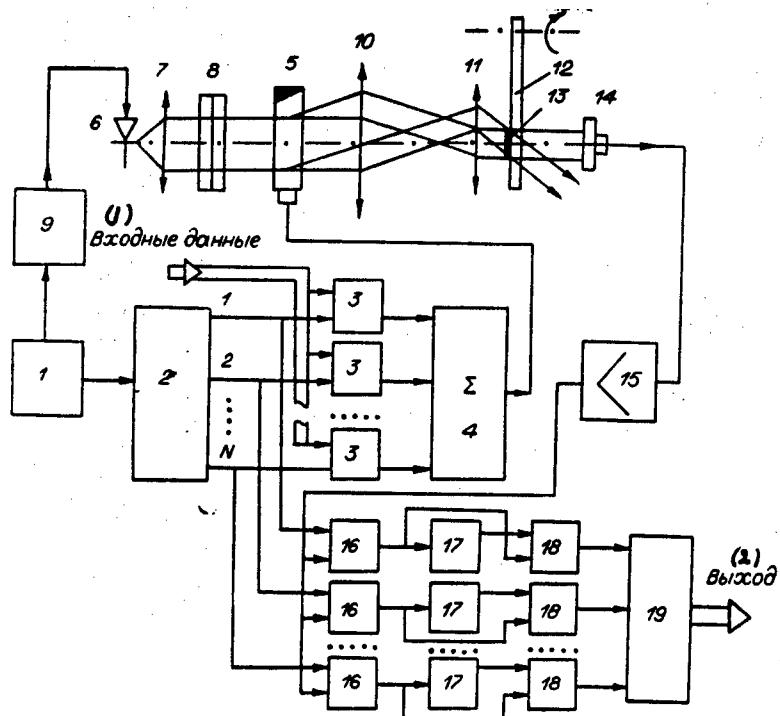


Figure 3.
Key: 1 — input data; 2 — output.

second. The energy of spurious phase fluctuations is concentrated in the range of space frequencies of 0-1 lines/mm and time frequencies 0-100 Hz.

The method of relative phase modulation helps reduce the effect of slow fluctuations of memory parameters and, therefore, a higher absolute stability is not required: these parameters have only to remain practically constant on the writing/reading interval of two neighboring holograms. This interval is equal to 1-10 μ s in time and 3-10 μ m in space, which is much less than the typical length and time of variation of memory parameters.

For decoding the read-out signal, the result of measurement of the phase of the i th harmonic of the hologram is sent from the output of the digital phase detector 16 to the input of the buffer memory 17 and to one of the inputs of the arithmetic device 18. The other input of the arithmetic device receives from the buffer memory the value of the phase of the same i th harmonic of the preceding hologram. The arithmetic device computes the phase increment and determines by using the code alphabet in decoding unit 19 the value of the respective digit (or group of digits) of the read-out binary word.

The main features of the optical head are the following: hologram dimensions = $105 \times 3 \mu^2$; hologram capacity = 64 bits; registration speed = 10^6

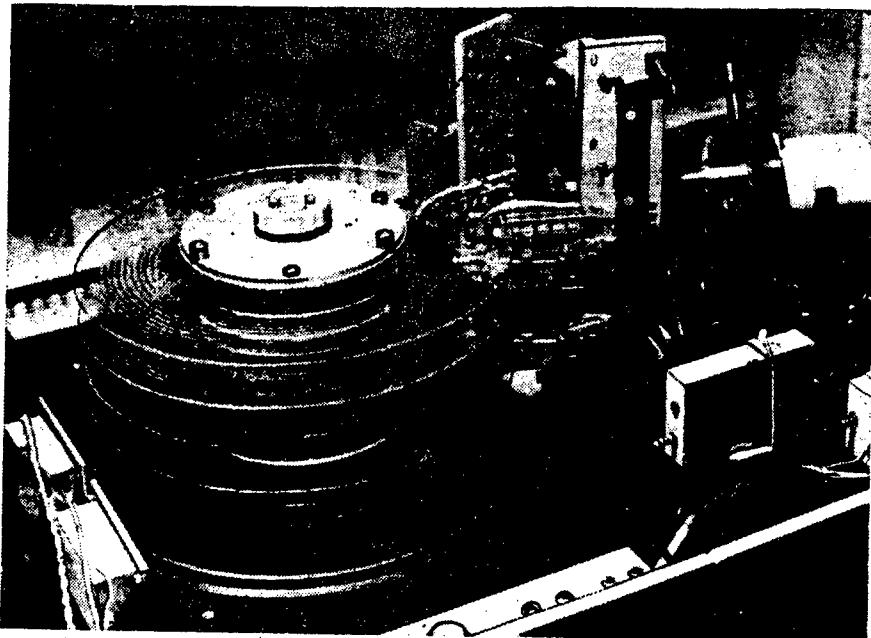


Figure 4.

holograms/s; information coding method = relative four-phase manipulation; reading speed = 1.25×10^5 holograms/s; head height (minimal distance between the disks in a pack) = 25 mm; minimal and maximal space frequencies of a Fourier hologram being formed = 680 and 1360 lines/mm; light efficiency = 10 percent; attenuation of disk beats by aerostatic suspension of the microlens = at least 500; pressure on the disk surface = 3.5-5 N; clearance between aerostatic suspension plate and disk surface = 30 μ m.

The parameters of the electric signal fed to the AOM are: frequency range = 64-128 MHz; frequency grid step = 2 MHz; signal power = ≈ 0.5 W. The semiconductor laser operates in continuous and pulsed modes; the emission power in continuous mode is at least 20 mW; the pulse energy for a duration of 4 ns (at one-half maximum intensity level) and maximum repetition frequency of 1 MHz = 4×10^{-9} J.

Figure 4 is the general view of the experimental memory pack. The number of disks in a pack can vary from 1 to 8. The capacity of a single working disk surface is 1Gb.

One-dimensional holograms are recorded on concentric tracks. For continuous registration, two drives of two optical head blocks are provided, which operate in succession; while one group of heads writes information, the other group is shifted to the next track. Each optical head writes and reads information from a single disk surface. The number of heads in a block can vary from 1 to 4.

Electronic blocks controlling the optical heads and electromechanical devices of the memory are mounted in two CAMAC crates; one of these is linked to a computer through a common bus interface. The computer controls the positioning of the write/read head, specifies disk pack rotation speed, determines writing and reading modes, and executes troubleshooting and data writing verification.

Methods and Results of an Experimental Study of the Memory Pack

In writing information with maximum density, one should take into account the frequency limitations that arise in the optoelectronic channel of the memory and can result in interference from neighboring digits of a word being recorded. In a linear approximation these distortions can be characterized by a transition function [PF].

The PF of an optoelectronic write/read system is one of the basic memory parameters. It is defined by the system response to a δ pulse — the signal read from a specially made diffraction grating much larger than the actual hologram. Such a diffraction grating that has a space frequency v_p is placed in the reading zone of the optical head (instead of the hologram). The voltage of linearly varying frequency [LChM] is fed to the electric input of the AOM from the spectral analyzer. The voltage from the output of the amplifier of the read-out signal is represented on the spectral analyzer screen.

During scanning, two crossing beams fall upon the reference diffraction grating: the (zero) reference beam, which has total amplitude $A_o(x)$, and the scanning beam, with complex amplitude $A_c(x)\exp(j\omega t)$, where $\omega = 2\pi f$ is the frequency of the voltage fed to the AOM piezotransducer. The optical head and the reference grating are oriented so that the vector of the reference grating and the grating formed by the crossing beams $A_o(x)$ and $A_c(x)$ are directed along the axis x , i.e., along the long side of the hologram. The variation rate of the frequency df/dt is taken such that within the AOM aperture the variation of the frequency of the ultrasound grating can be disregarded: $(df/dt)t_A^2 \ll 1$, where t_A is the AOM aperture time.

We will represent the complex amplitudes of the reference and scanning beams in the reference grating plane as a set of plane waves whose amplitudes and phases are determined by Fourier transforms (F):

$$S_o(v) = F[A_o(x)], \quad S_c(v) = F[A_c(x)],$$

where v is space frequency ($v = 2\pi/\Lambda$, Λ is the grating period).

Since the scanning is formed by diffraction on the ultrasound phase grating moving in AOM, we can write, in the absence of amplitude and phase

distortions, for the scanning beam immediately behind the reference grating plane: $S_c(v) = kS_o(v + n_1\omega/V - \nu_p)\exp(j\omega t)$. Here, $S_o(v) = S_o(v)\exp[j\varphi(v)]$; k is the coefficient which depends on the diffraction efficiencies of AOM and the grating; V is the speed of sound in AOM; $n_1 > 1$ is the reduction coefficient of the telescopic system of the optical head.

We denote by $\exp[jQ(v, z)]$ the phase characteristic of the medium situated between the reference grating and the photodetector (at the distance z). The photodetector is installed in the reference (zero) beam, and its size is such that it fully covers the beam. The amplitudes of the reference beam and the scanning beam diffracted on the grating are combined on the photodetector. Accordingly, for the variable component of the photocurrent we can write

$$\begin{aligned} I_f(\omega, t) &= K_f \int_{-\infty}^{+\infty} [S_c(\nu) + S_o(\nu)] \exp[jQ(\nu, z)][S_c(\nu) + S_o(\nu)]^* \exp[-jQ(\nu, z)] d\nu \\ &= K_f K_1 \int_{-\infty}^{+\infty} S_o(\nu) S_o(\nu - \nu_p + n_1\omega/V) \cos[\omega t + \varphi(\nu - \nu_p + n_1\omega/V) - \varphi(\nu)] d\nu. \end{aligned}$$

Further, considering that

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} S_o(\nu) S_o^*(\nu - \nu_p + n_1\omega/V) d\nu = \int_{-\infty}^{+\infty} |A_o(x)|^2 \exp[-f(n_1\omega/V - \nu_p)x] dx,$$

we obtain an expression for the photocurrent amplitude:

$$I_{FA}(\omega) = K_2 \left| \int_{-\infty}^{+\infty} P(x) \exp[-j(n_1\omega/V - \nu_p)x] dx \right|,$$

where $P(x) = A_o^2(x)$ is beam intensity distribution at the AOM input; K_f , K_1 and K_2 are coefficients. Thus, $I_{FA}(\omega)$ is virtually the "bit image" in the plane of space frequencies v . Accordingly, $I_{FA}(\omega)$ has the same meaning for a Fourier hologram as does the point scattering function for bit-by-bit registration devices.

From the form of PF one can evaluate quantitatively the accuracy of the setting of the microlens focal plane. If with the air suspension system activated the diffraction grating plane does not coincide with the microlens focal plane, PF becomes asymmetric relative to the ordinate axis and side lobes are formed to the right or left of it.

PF in fig. 5a represents the case where the microlens focal plane coincides with the grating plane. The PF maximum is at 107 MHz. The oscillogram scale on the abscissa is with one division = 0.5 MHz or 5.3 lines/mm for the space frequency region (the frequency range of the hologram is 680-1360 lines/mm). A PF with an improperly placed microlens plane is shown in fig. 5b; the

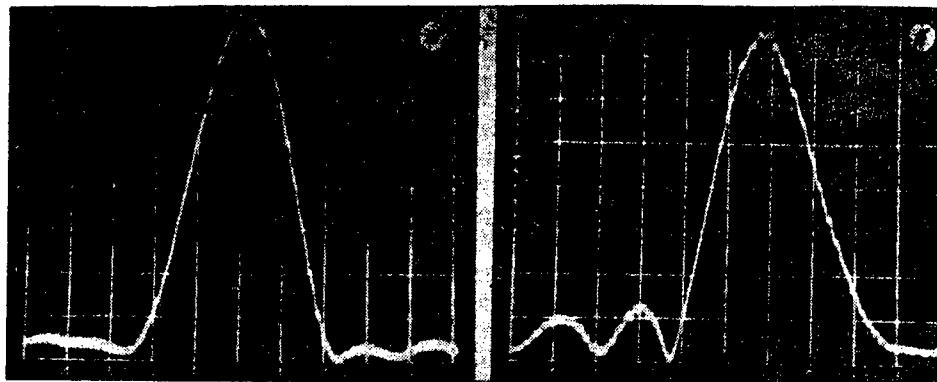


Figure 5.

grating is approximately 25 μm behind the focal plane. In case of opposite detuning, side lobes would appear to the right of PF.

An inaccurate setting of the focal plane makes the PF maximum dependent on the movement of the microlens along the optical axis because of disk wobble. A PF shift reduces the level of the signal read and increases intercharacter noise. To simulate disk wobble the diffraction grating was moved along the optical axis ± 0.15 mm. No significant (>20 percent) widening of PF was observed; the central frequency shifted at most ± 200 kHz.

The PF width depends largely on the adjustment of the optical head and the width of the direction pattern of laser radiation in the plane perpendicular to the p-n-junction plane. From known head dimensions (aperture time of AOM and truncation level of the Gaussian beam) one can calculate PF width and, comparing it with the experimental width, estimate qualitatively the level of optical aberration. For instance, the PF in fig. 5 was obtained with the following head parameters: beam size at AOM input in the direction of the sound propagation = 0.806 mm; beam truncation level = 0.5. Since the speed of sound in the light-soundguide material of the modulator is 0.72×10^3 m/s, the estimate of aperture time $t_A = 1.12 \mu\text{s}$. For a truncation level of 0.5, calculated PF width at the 0.5 level is 1.16 MHz. From a comparison with experimental PF width (1.25 MHz) one can characterize the resolution power of the head, which is close to the diffraction limit.

Figure 6 presents a fragment (24 of 32 bits) of the signal read during continuous scanning of the frequency of AOM control voltage. The bits of the written word are reproduced on the spectral analyzer screen successively. The ninth bit (82 MHz) is recorded with amplitude coding and is equal to zero. The scale on the abscissa is one division per 5 MHz. From an analysis of the image the main signal characteristics can be evaluated: the voltage level of logical one (log.1), logical zero (log.0) and especially the displacement of extremum points of the image relative to the discrete grid of

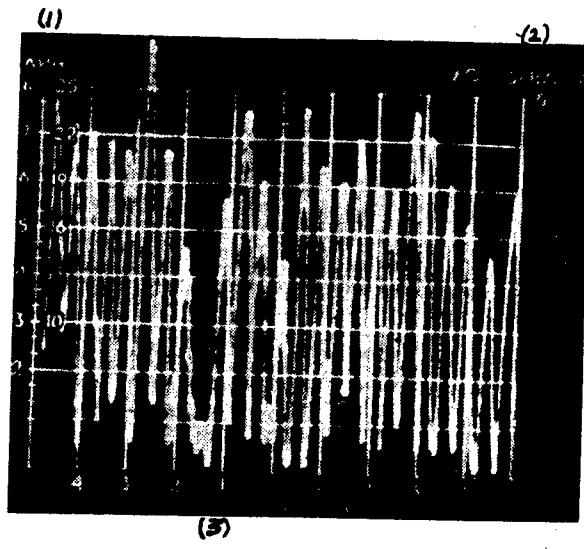


Figure 6.
Key: 1 — linear; 2 — log
division; 3 — central frequency.

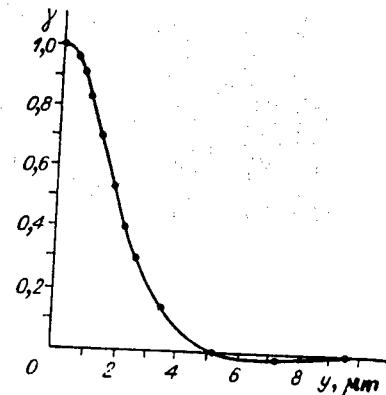


Figure 7.

synthesizer frequencies (64, 68, ..., 100, ..., 128 MHz). The main cause of the displacement is the setting of emulsion during photochemical treatment. The maximum is shifted relative to the synthesizer frequency ± 0.3 MHz (± 0.23 percent at boundary frequency of 128 MHz); this reduces the response level by approximately 10 percent [1].

Two photoemulsion treatment methods were investigated: SGZh bleaching [haloid silver gel] technology [6] and clarification. The only acceptable method for recording phase holograms on disks was clarification, which provided a frequency shift of individual hologram gratings by at most ± 0.28 MHz.

Another important parameter of the optical head is the effective width of the hologram: the transverse size of the interference grating formed at the intersection in the "image plane" (see fig. 1) of the zero (reference) and the diffracted beams. Aberrations of the optical scheme result in an imprecise combination of these beams in the transverse direction, reducing the effective width and the contrastivity of the hologram and lowering the signal-to-noise ratio.

The effective width of the hologram was determined from the autocorrelation function of the speckle signal read from a specially made disk with an exposed and bleached photoemulsion. The autocorrelation function width depends on the linear speed of the disk and the time of variation of the random surface of the photographic medium within the hologram width.

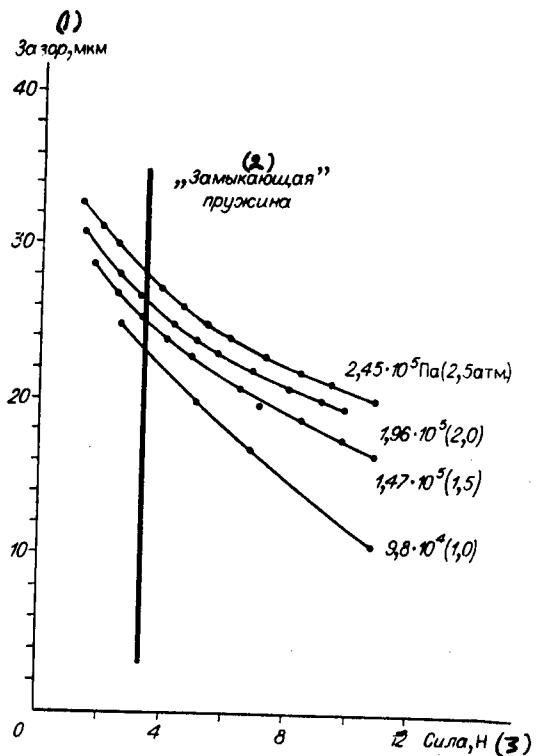


Figure 8.

Key: 1 — clearance, μm ; 2 — "locking" spring; 3 — force, N.

A typical form of experimental autocorrelation function $\gamma(y)$ of read-out noise of photographic medium scattering is illustrated by fig. 7. We see that as the hologram is shifted along the track to $y \approx 5 \mu\text{m}$ the speckle noise that is read becomes virtually uncorrelated.

If the Gaussian beam has a stricture width $W = 3 \mu\text{m}$ (at the half-maximum intensity level), then with admissible widening of W to 10 percent the variation of the distance Δ to the focusing plane should be at most $\Delta \approx W^2/\lambda \approx 10 \mu\text{m}$, where $\lambda = 0.85 \mu\text{m}$ is the wavelength of light.

Figure 8 plots the clearance between the aerostatic suspension plate and the disk as a function of the force pressing the plate to disk surface. The characteristics were obtained with various positive pressures in the air pipeline feeding the aerostatic suspension. The "locking" spring has initial force of 3.5 N and rigidity $C_p = 0.5 \text{ N/mm}$. The intersection points of the characteristics of the mechanical and aerostatic springs determine the height of the "hovering" and the rigidity of the "spring" C_v of the microlens air suspension. With a positive pressure in the pipeline of $2.45 \times 10^5 \text{ Pa}$ (2.5 atm) these values are $28 \mu\text{m}$ and 0.5 N/mm , respectively.

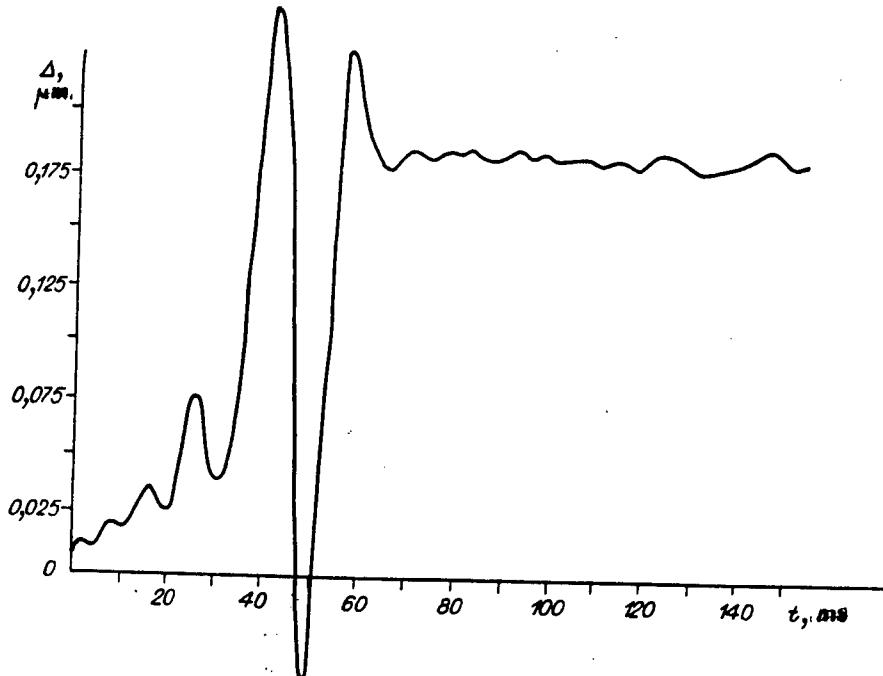


Figure 9.

The degree of damping of wobbles, i.e., the ratio of disk wobble along the pack axis ξ to the variation of the spacing Δ between disk surfaces and the suspension plate, is defined by the expression

$$\varepsilon(\omega_d) = \frac{\xi}{\Delta} = \left[\frac{(\omega_{0p}^2 + \omega_{0v}^2 - \omega_d^2)^2 + \omega_d^2(H_p + H_v)^2/m^2}{(\omega_{0p}^2 - \omega_d^2)^2 + \omega_d^2H_p^2/m^2} \right]^{1/2}, \quad (1)$$

where m is the mass of the moving component of the aerostatic suspension; ω_d is the circular disk wobble frequency; $\omega_{0p} = \sqrt{C_p/m}$, $\omega_{0v} = \sqrt{C_v/m}$; H_p and H_v are resonance circular frequencies and viscous friction coefficients of the spring suspension and the air suspension, respectively.

The mass of the moving suspension component near the optical head (see fig. 2) is $m = 50$ g. Hence, $\omega_{0p} = (2\pi \times 16)$ 1/s; $\omega_{0v} = (2\pi \times 500)$ 1/s. The maximum disk rotation speed is 4 rps; therefore, in the working frequency range ε is largest at $\omega_d = 0$; it is $\varepsilon(0) \approx 1000$.

Figure 9 shows experimental $\Delta(t)$ as ξ is changed in a jump to 100 μm . The frequency band of the measurement device is 0-100 Hz. Figure 10 shows the frequency characteristic $\varepsilon(\omega_d)$ calculated from $\Delta(t)$.

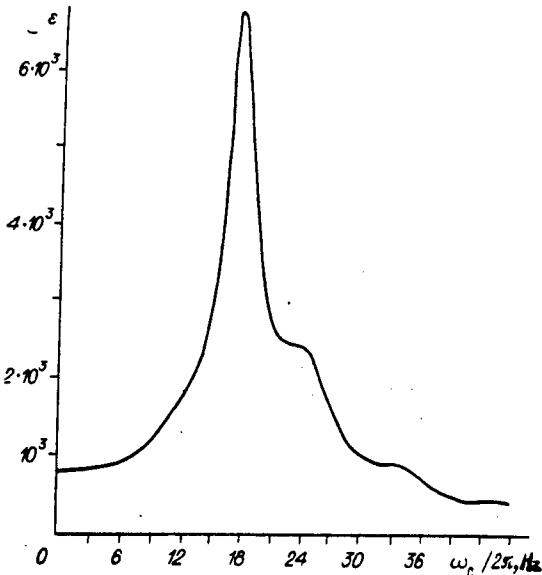


Figure 10.

Reliability Analysis

All additive noises operating in the reading/writing system can be subdivided into two basic groups: one group includes the noise whose individual realizations are time functions; the second group consists of noise determined by the hologram structure and described by functions of the space coordinate.

The first group includes thermal and shot noise of electronic devices and redundant noise of semiconductor laser radiation. For reducing such noise one should narrow the passband of the reading device, i.e., ultimately reduce information reading speed.

In the second group, optical scattering noise of the head and the registration material are most important. The level of this noise depends on the space frequency transmittance band of the optical scheme; reducing the passband in this case will reduce the density of recording, which is the most important parameter of an optical memory.

We will examine the influence of the main sources of time noise on the signal-to-noise ratio [S/Sh]. If the reading beam power is limited [7],

$$(S/N) = \frac{I_{f_0} \kappa_g [2\eta_m \eta_g (1 - \eta_m) t_n]^{1/2}}{\kappa_m N [2eI_{f_0} (1 - \eta_m) + \beta I_{f_0}^2 (1 - \eta_m)^2 + 2eI_{sh}]^{1/2}}, \quad (2)$$

where η_m , κ_m ; $\eta_g \ll 1$, κ_g is the diffraction efficiency and light transmittance in amplitude of the AOM and the hologram, respectively; N is the number of bits in the hologram; I_{sh} is the photocurrent, whose shot noise is equal to the amplifier noise; t_n is accumulation time; e is the charge of an electron; $I_{f_0} = P_c K_c$ is the constant of the photodiode current at $\eta_m = 0$, $\kappa_g = 1$; P_c is the intensity of the reading beam at the memory output; K_c is the transformation coefficient of the reading head. According to (2), an increase of I_{f_0} will cause the signal-to-noise ratio to approach the limit:

$$\max(S/N) = \frac{\kappa_g [2\eta_m \eta_g t_n]^{1/2}}{\kappa_m N [\beta(1 - \eta_m)]^{1/2}}, \quad (3)$$

Estimate (3) holds if we can disregard the shot noise of the photocurrent $\sim 2eI_{f_0}(1 - \eta_m)$ and the amplifier noise $\sim 2eI_{sh}$ as compared with the excess noise of the semiconductor laser $\sim I_{f_0}^2 \beta(1 - \eta_m)^2$. The first condition is formulated as $I_{f_{01}} \gg 2e/\beta(1 - \eta_m)$; the latter condition is

$$I_{f_{02}} > [2eI_{sh}/\beta]^{1/2}/(1 - \eta_m). \quad (4)$$

The photocurrent amplifier in the working frequency band of 66-128 MHz has input-reference noise which is equivalent to the shot noise on the order of 200 μ A ($I_{sh} = 200 \mu$ A). In the best lasers that have been studied, $\beta = 2 \times 10^{-14}$ s; accordingly, for $\eta_m = 0.2$ the S/N ratio is close to the maximum (3) at $I_{f_{01}} \gg 20 \mu$ A, $I_{f_{02}} > 71 \mu$ A.

The power of the reading beam and the head output $P_c = 2.5$ mW; the coefficient $K_c = 0.08$ mA/mW; $I_{f_0} = K_c P_c = 200 \mu$ A. Accordingly, in first approximation, conditions (4) are satisfied. In that case, if $\kappa_g = 0.9$, $\eta_g = 0.01$, $\kappa_m = 0.9$, $N = 32$, $t_n = 1 \mu$ s (reading speed 64Mb/s), then $S/N \approx 15$. In the absence of redundant laser noise, $S/N = 33$.

This example shows that even in the best lasers the excess noise is so great that it has a drastic effect on the signal-to-noise ratio. In many lasers β is as large as 10^{-13} s. One has to increase the accumulation time t_n , reducing reading speed. In the memory model t_n was set at 5.4 μ s. Considering measurement and storage clearance time, the cycle of reading a single hologram is 8 μ s; the reading speed is 8Mb/s.

The scattering noise of the optical elements does not affect the result of information reading, because heterodyne reading occurs only in the zone with moving diffraction gratings formed at the intersection of the reference (zero) beam and the beam diffracted in AOM. The optical scheme of the head is such that these beams overlap only on the disk surface in the zone with the hologram. Another important factor is that at the AOM output the diffracted beam and the zero beam are orthogonally polarized and cannot form an interference grating. In experiments it was determined that in sequential

reading (by a scanning beam) the head detects reliably the scattering noise of the disk material (thermally polished glass).

Photoemulsion scattering noise affects registration reliability to the greatest degree [1]. The spectral density of this noise after emulsion treatment with clarification is at the level of $\Phi_{sh} = 10^{-8} \text{ mm}^2$. The effective frequency band of the reading beam, which is 105 μm long and 3 μm wide (at half-maximum intensity level), is $B = 2200 \text{ 1/mm}^2$. The diffraction efficiency of noise lattices $\eta_{sh} = \Phi_{sh}B = 2.2 \times 10^{-5}$ and $S/N = 14$ (at $\eta_s = 0.01$, $N = 32$). Considering excessive laser radiation noise and the noise from variation of hologram diffraction efficiency, depending on track number, we obtain S/N from 7.3 to 11.3. Each value results from an averaging in a file of 1000 holograms.

The probability density for the phase $p(\varphi)$ of the sum of a narrowband Gaussian stationary process and deterministic harmonic signal [8] at large $S/N = a > 5$ can be expressed as

$$p(\varphi) \simeq \frac{a}{\sqrt{2\pi}} \cos \varphi \exp\left(-\frac{a^2}{2} \sin^2 \varphi\right). \quad (5)$$

The resulting angle ($\Delta\varphi$) is determined by the phase shift of equal space frequency of two neighboring holograms distorted by an uncorrelated noise component. The probability density of the resulting angle was determined numerically as the convolution of the probability densities $p(\varphi)$ (5). For relatively high signal-to-noise ratios the losses from the use of this correlation detection instead of coherent detection [5] was approximately 3 dB. Nevertheless, the memory uses correlation detection because for multichannel parallel reading systems the detection device is thus greatly simplified.

Figure 11 is the histogram of distribution of phase difference based on 4096 measurements. The ratio of signal amplitude to its standard deviation for this diagram is 10.2.

Table 1 gives for 10 files (4096 measurements each) the number of overshoots n_4 and n_8 beyond threshold levels, equal to $\pm\pi/8$ (four-phase modulation) and $\pm\pi/2$ (eight-phase modulation), respectively. There were no overshoots beyond the $\pm\pi/2$ level. The column n_A indicates for these 10 files the number of errors for amplitude information coding in the hologram. For these files there are no data on the distribution of log.0 levels; only errors of the "1→0 transition" are indicated. The threshold of a_p/a_c for each file is set near the optimal value [1].

The theoretical (calculated) estimate of the number of deviations for file 2 is $n_4 = 1$ and $n_8 = 262$; for file 7, $n_8 = 32$.

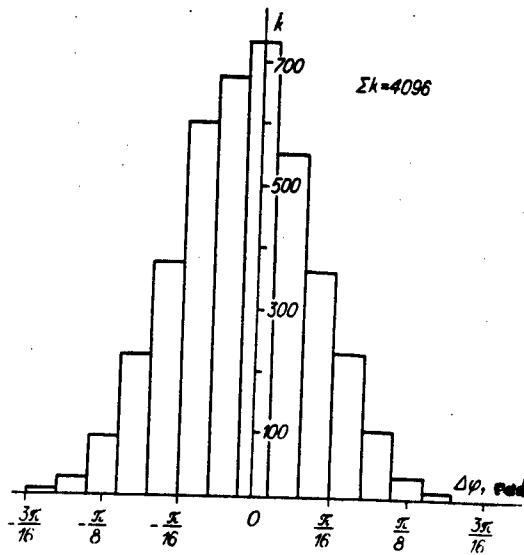


Figure 11.

Table 1.

File number	S/N	σ , rad	$n_4(\pm\pi/4)$	$n_5(\pm\pi/8)$	a_p/a_c	$n_A(1 \rightarrow 0)$
1	7.3	0.22	4	300	0.56	2
2	7.3	0.22	2	285	0.56	15
3	7.8	0.18	-	144	0.54	2
4	7.8	0.18	1	142	0.54	4
5	7.9	0.17	-	115	0.54	2
6	8.5	0.17	-	135	0.53	-
7	10.2	0.176	-	111	0.53	2
8	10.2	0.15	-	41	0.53	-
9	10.5	0.16	-	54	0.53	-
10	11.5	0.16	-	66	0.53	-

From a comparison of n_4 and n_A it follows that the four-phase coding combined with correlation detection provides noise immunity not inferior to that with amplitude coding. The value of a_p/a_c in the table for each data file was optimal, but the practical implementation of this method for a 32-channel device involves considerable technical difficulties. By contrast, phase demodulator detectors are triggered at zero threshold, which is the same for

all channels. Besides, four-phase coding increases registration speed and density by a factor of 2.

The phase shift value can be converted rather easily to a digital code: this offers an opportunity for improving registration reliability by combining multilevel (eight-phase) modulation and a convolution code with rate $\frac{2}{3}$ and a nonrigid decision-making algorithm.

Conclusions

A model of optical disk pack memory has been built which interacts with a computer and is to be used for studying a variety of digital data writing/reading techniques.

A small-sized optical write/read head and electronic control units have been designed and studied experimentally; they implement four-level phase information coding in holograms. The head provides a writing speed of 64Mb/s, registration density of 2×10^5 bit/mm² and a reading speed of 8Mb/s.

This study showed that the method of relative (difference) phase modulation reduces the influence of slow fluctuation of memory parameters upon the reliability of data registration, makes the requirements for identity of parallel channels less stringent and thus creates conditions for organizing multilevel information coding in holograms.

With the methods proposed here, it is possible to estimate the basic characteristics of the memory device: the transition function of the optoelectronic data writing/reading channel, the resolution of the optical head, the level of intercharacter interference and other parameters of a signal read out of a hologram.

The main source of additive noise is excess noise of the semiconductor laser and the scattering noise of registration material. The high level of excess laser noise does not allow developing a reading speed equal to data registration speeds. A better reliability could be achieved not only by improving the parameters of the medium and its treatment conditions, but also by means of state-of-the-art methods of coding and error detection and correction.

Bibliography

1. Vovk, Yu.V., et al., "Method of Multichannel Registration of Binary Data on an Optical Disk," AVTOMETRIYA, No 2, 1989.

2. Vovk, Yu.V., and Shchepetkin, Yu.A., "Frequency Signal Separation Used for Registration of One-Dimensional Holograms by Semiconductor Laser Radiation," AVTOMETRIYA, No 1, 1981.
3. Vovk, Yu.V., and Shchepetkin, Yu.A., "Parallel Heterodyne Reading of Binary Data on One-Dimensional Fourier Holograms," AVTOMETRIYA, No 3, 1984.
4. Trubetskoy, A.V., and Shipov, P.V., "Multifrequency Anisotropic Modulators on TeO₂ Crystals," AVTOMETRIYA, No 3, 1989.
5. Zayezdnyy, A.M., Okunev, Yu.B., and Rakhovich, L.M., FAZORAZNOSTNAYA MODULYATSIYA I EYE PRIMENENIYE DLYA PEREDACHI DISKRETNOY INFORMATSII [Phase Difference Modulation and Its Use in Discrete Information Transmission], Moscow, Svyaz, 1967.
6. Mikhaylov, I.A., and Vanin, V.A., "Holograms on Haloid Silver Gel," in: SHESTAYA VSESOYUZNAYA SHKOLA-SEMINAR PO OPTICHESKOY OBRABOTKE INFORMATSII: TEZISY DOKLADOV [Sixth All-Union School-Seminar in Optical Information Processing: Abstracts of Papers], Frunze, FPI, 1986.
7. Zatolokin, V.N., Shteynberg, I.Sh., and Shchepetkin, Yu.A., "Influence of Semiconductor Laser Noise in Heterodyne Information Reading from Holograms," in: III VSESOYUZNAYA KONFERENSIYA PO VYCHISLITELNOY OPTOELEKTRONIKE "PROBLEMY OPTICHESKOY PAMYATI": TEZISY DOKLADOV [Third All-Union Conference on Computational Optoelectronics, "Optical Memory Problems": Abstracts of Papers], Yerevan, Izdatelstvo ArmSSR, 1987.
8. Goryainov, V.T., Zhuravlev, A.G., and Tikhonov, V.I., STATISTICHESKAYA RADIOTEKHNIKA: PRIMERY I ZADACHI [Statistical Radioengineering: Examples and Problems], Moscow, Sovetskoye radio, 1980.

COPYRIGHT: "Nauka," "Avtometriya," 1989

UDC 681.327.664.4

Bubble Memories

907G0030A Kiev UPRAVLYAYUSHCHIYE SISTEMY I MASHINY in Russian No 5, Sep-Oct 89
(manuscript received 22 Feb 88, after correction 17 Jun 88) pp 43-47

[Article by N. B. Malinovskiy under the rubric "Microprocessor Engineering and Microtechnology"]

[Text] Introduction

Bubble memories, representing a new trend in memory engineering, have received development both in our country and abroad in the last decade. Research and development in this field were begun in 1967 by Bell Telephone Labs in the USA, and then IBM, Rockwell International, Fujitsu Labs, etc., joined in this research.

A coalition of French industrial firms, as well as Intel, Hitachi, Ltd., Fujitsu, Ltd., and Plessey Microsystems, Ltd., have now taken the leading position in the development of the bubble technology abroad [1-3].

Bubble memory systems differ advantageously from disk and tape storages by the absence of mechanical units and have a whole series of advantages such as high reliability, resistance to mechanical, climatic and special influences, the preservation of data when the power is cut off, high information density, low power requirement, and relatively high speed, although they are inferior with respect to cost and capacity.

The greatest interest in the bubble technology is being displayed by military customers, manufacturers of aerospace equipment and numerically controlled machine tools and other consumers needing highly reliable large-capacity nonvolatile memories [2, 4, 5].

Bubble Microcircuits

A typical bubble microcircuit chip consists of a substrate (e.g., of gadolinium gallium garnet) on which an epitaxial film of a magnetic material is grown. This film is the medium in which bubbles exist under the influence of an external magnetic field [6]. The bubbles' diameter can equal from tenths of a micrometer to full micrometers depending on the type of magnetic material.

Functional control units--a generator, annihilator, input/output switches--as well as ferromagnetic applications for producing moving circuits (registers) and bubble reading sensors, are formed by thin-film technology and lithography.

The generator serves the function of producing bubbles, and the annihilator of erasing them. Bubbles are divided (replicated) into two individual bubbles by means of input/output switches, and information is shifted and stored in storage registers (the presence of a bubble represents logical "1", and the absence of one, logical "0").

Control functions are executed by the supplying of current pulses to individual control elements. The current passing through the control element causes a local change in the magnetic field. This change in the magnetic field together with the moving field produced by two orthogonal coils performs control functions. Bubbles can be attracted and repelled, sending them from one register to another, by means of the magnetic field created by control elements. This forms the basis for the implementation of the transfer function.

There are two main types of bubble microcircuits: of the serial type, based on one open or closed storage register, and of the serial-parallel type, in which from dozens to several hundred storage registers function simultaneously, united by means of input/output registers with the serial input and output of the data sequence.

The type FBM31DB bubble microcircuit made by Fujitsu can be cited as a typical example of a bubble memory based on an off-line closed storage register. This microcircuit's capacity is 74,032 bits, its average access time is 370 ms, its data exchange rate is 100K bits/s, and power requirement 0.5 W (the data are taken from the company's brochure). A bubble generator, bubble annihilator, replicator, off-line closed storage register, and two working and two compensating reading sensors are placed on the chip of this microcircuit.

Microcircuits of a similar type have been developed also by Rockwell International, with a capacity of 102,400 bits, Bell Telephone Labs, with a capacity of 68,121 bits, and Plessey Microsystems, Ltd., with a capacity of 69,712 bits. A domestic development is a microassembly having a capacity of 4225 bits [7].

Increasing the capacity of bubble microcircuits having a serial structure entails a corresponding increase in access time. Therefore, they have not become widely used and have not been developed further, but they were made at the initial stage of the development of the bubble technology.

An important step in the development of bubble memories was the change to the serial-parallel organization of the structure of a bubble microcircuit. This organization makes possible a much shorter information access time than that of a purely serial organization, and it also makes it possible to introduce redundancy for the purpose of increasing the yield of acceptable bubble microcircuits. Besides, by improving the technology and methods of

fabricating bubble microcircuits it is possible to achieve an increase in information recording density without a substantial increase in access time.

Typical bubble microcircuits having a serial-parallel organization are the type RBM 256 microassembly having a capacity of 256K bits from Rockwell International [8], the IM7110 and IM7114 having a capacity of 1M and 4M bits, respectively, from Intel Magnetics [9, 10], and the domestically produced K1605RTs1 and K1602RTs2 having a capacity of 256K bits, and the K1602RTs3 having a capacity of 1M bits [10-12], etc.

A generator, input and output switches, two working and two compensating reading sensors, input and output registers, and 282 information storage registers are placed on a chip of the RBM 256 microassembly, for example, having an area of 1 cm². Of the 282 storage registers, 22 registers can be rejected. The yield of acceptable bubble chips is increased because of this redundancy. The numbers of the rejected registers are recorded in a ROM, for example, and represent what is termed the suitability chart of storage registers. The writing of information to rejected registers is not permitted.

The capacity of each storage register is 1025 bits, i.e., the microassembly's total useful capacity is $1025 \times 260 = 266,500$ bits. As a rule, 256K bits are used in practice for the storage of information, and the remaining space can be filled with auxiliary information.

The average access time of the RBM 256 microassembly is 4 ms with a control field frequency of 150 kHz, and 6 ms with a frequency of 100 kHz. The microcircuit consumes approximately 1 W in the operating mode. Its operating temperature range is from -10 °C to +65 °C at a frequency of 150 kHz, and from -10 °C to +70 °C at a frequency of 100 kHz. Its temperature range in the storage mode is from -50 °C to +100 °C.

The properties of the type K1605RTs1 microassembly are similar to those of the RBM 256 [10, 11].

The most advanced bubble microcircuits now from the viewpoint of design and technology are microcircuits having a capacity of 4M bits. The further efforts of developers of bubble microcircuits are directed at expanding the operating temperature range to from -55 °C to +125 °C, increasing the data transfer rate, and increasing the capacity of bubble chips to 64M bits and more [3].

Bubble Movement Control

All bubble microcircuits having a capacity of up to 4M bits are now made with control coils and access by means of a rotating magnetic field (field access). The main shortcoming of these bubble microcircuits is the need to create a high-frequency rotating magnetic field. It has been established theoretically that bubble microcircuits can operate potentially with a frequency of up to 500 kHz, but, practically speaking, the upper limit does not exceed 200 kHz [9]. This limitation is due to the power dissipated on account of losses

resulting from the skin effect in the control coils and losses in the case's metal parts from eddy currents.

The rotating field in bubble microcircuits is usually excited by a pair of orthogonal coils (an inner and outer). The coils' dimensions are chosen to be sufficiently great in order to make possible a uniform rotating field over the entire area of the chip. In an effort to reduce overall size and dissipated power, Japanese developers have proposed a new case with a core in the form of a loop [3,13]. X- and Y-coils are wound on the parallel sides of the core and they are put into a shield made of a current-conducting material. The core's scattered field performs the control function. At frequencies of the rotating field on the order of hundreds of kilohertz the surface currents in the shield prevent the magnetic field from going beyond the shield's bounds. This fact makes possible a more uniform magnetic field inside the shield. The shielding effect is frequency dependent and is most effective at frequencies above 5 kHz. The use of a case of this design makes it possible to reduce the size of a bubble microcircuit approximately fourfold as compared with traditional designs.

Increasing the chip's capacity when field access is used entails the problem of lengthening the access time. One method of solving this problem is based on the use of current access. Bell Telephone Labs has proposed a technology for bubble microcircuits that is based on the use of two conducting layers [8, 14]. Bubbles can be advanced by controlling each conducting layer by means of four-phase current pulses. This technology could solve the problem of lowering the supply voltage to 5 V (supply voltages of from 12 to 20 V are required for field-access bubble microcircuits) and of improving performance, as well as of doing away with the use of control coils. Nevertheless, because of their great dissipated power, current-access devices have not as yet acquired a practical application.

A Z-shaped coil placed in the chip's plane [5] is added in some recent bubble microcircuits for the purpose of expanding their temperature range. It reduces the influence of external magnetic fields on the memory's magnetic field and prevents bubbles from collapsing at elevated temperatures. The instantaneous erasure of information can also be carried out by means of this coil. This is very important for a number of special-purpose memory systems.

Promising Bubble Technologies

Discrete, as a rule, permalloy applications are the moving structures in present-day bubble microcircuits. A recording density of up to $1.5M \text{ bits/cm}^2$ can be obtained as a maximum in bubble microcircuits of this generation by means of the available lithography technique [15, 16].

The generation of bubble microcircuits having a capacity of from 4M to 16M bits will most likely be fabricated by the use of moving structures in the form of adjoining disks [9, 15], which can be formed by means of ion implantation or permalloy evaporation. The increase in packaging density in devices of this sort is associated with the fact that the permissible resolution of lithography can exceed a bubble's diameter by a factor of 1.5 to 2, which makes it possible at lithography's present-day level to make a bubble structure having bubbles $1 \mu\text{m}$ in diameter. Bell Telephone Labs researchers

were among the first to create an experimental chip measuring 28 x 30 mm having a structure of the "adjoining disk" type, that is able to store as a maximum 11,542,272 bits [17].

There are at the present time neither exchange gates nor block replicator gates that have been implemented in practice that use ion-implanted structures. Therefore, hybrid structures in which the storage array consists of ion-implanted structures and the input and output registers consist of permalloy structures [13] are used more in practice. The main complication in making hybrid structures consists in joining the ion-implanted structure to the permalloy structure, because it is necessary to provide for the advancement of bubbles from ion-implanted to non-ion-implanted regions, i.e., to cross the potential barrier.

The next, completely new stage in the development of bubble memory technology is a Bloch line memory [3, 13]. It has been established that the domain boundary of cylindrical magnetic domains contains vertical Bloch lines. They are suppressed in ordinary bubble microcircuits in order for the correct motion of the bubbles not to be disturbed. At the same time, pairs of Bloch lines can be used for the purpose of representing binary data. When the bias field is changed, the domain boundary moves and a pair of Bloch lines moves along the boundary under the influence of the gyrotropic force. A local Bloch loop that results in the formation of a pair of Bloch lines is formed when the tip of a strip domain is excited. Writing can be carried out in this manner.

A Bloch line memory makes it possible to obtain information densities on the order of hundreds of megabytes per square centimeter and to accomplish a very rapid parallel associative search for information in giant arrays. Whereas in ordinary bubble microcircuits one domain represents one bit of data, in Bloch line devices one strip domain 0.5 μm wide can store up to 100 bits [18]. Research is now under way in this area in the USA, Japan and the FRG.

Control Circuits

Regardless of the specific design of a bubble microcircuit, a typical finished bubble memory system will always contain on a board or in a cassette one or more bubble microcircuit cases supplemented with control and reading circuits, i.e., bubble microcircuits must be supplemented with special-purpose servicing electronics LSI circuits in order to construct finished miniature memory systems having high technical performance. As a rule, these LSI circuits replace from 20 to 50 cases of ordinary integrated circuits [19, 20].

Intel Magnetics has developed an LSI chip set for the type 7110 bubble microassembly having a nominal capacity of 1,048,576 bits. The auxiliary LSI circuits of this set, in providing a simple interface, at the same time retain the flexibility the user needs in designing memory systems according to his requirements. The set of auxiliary LSI circuits includes a type 7220 bubble memory controller, a formatting device--a type 7242 reading amplifier, a type 7250 control coil current predriver, and a type 7230 current pulse generator [21].

National Semiconductor Corp. developed a similar set of LSI chips for the type NBM 2256 bubble microassembly having a capacity of 256K bits, containing the following: the type INS82851 controller, DS3615 control current function driver, DS3616 coil current driver and DS3617 reading amplifier [22].

Domestic industry will commence production of the KM1144 series set of multi-purpose control circuits, including the KM1144AP1 movement current driver, the KM1144UL1 reading amplifier and the KM1144AP2 control current function driver [12]. This set is designed for controlling series K1602RTs2, K1602RTs3 and K1605RTs1 microcircuits. The type K1806VP1-103 and K1806VP1-157 microcircuits have also been developed, designed for forming a time sequence of control signals for bubble microcircuits of the K1602RTs2 (K1605RTs1) type [23]. The K1806VP1-103 microcircuit has a free interface, which makes it possible to use it in equipment having various architectures. The K1802VP1-157 microcircuit has a byte bus and can be used as part of microprocessor systems with addressing corresponding to an input/output port or memory cells in the microprocessor's address space.

Bubble Memory Systems

All kinds of memory systems are being developed on the basis of bubble microcircuits. They can be external memories for microcomputers, personal computers and terminals, all kinds of on-board storages, automatic answering devices, etc. [10].

For example, the bubble cassette memory system called Bublset developed by National Semiconductor clearly demonstrates the capabilities and merits of bubble memory systems [26, 25]. This subsystem has dimensions of 65 x 100 x x 140 mm, and the removable cassette, which is a plug-in nonvolatile solid-state data medium having a capacity of 256K bits and 1M bits, has dimensions of 46 x 51 x 22 mm. A feature of the system is its broad temperature range of from -20 °C to +70 °C. The Bublset subsystem's main characteristics are as follows: supply voltage--+5 V, +12 V, power requirement--4 W, average speed--8K bits/s, and maximum access time--30 ms for cassettes having a capacity of 256K bits, and 60 ms for cassettes having a capacity of 1M bits.

The bubble microcircuits operate with an operating cycle on-off time ratio of 10 percent in this system in order to raise the maximum permissible operating temperature. This small on-off time ratio lowers to a few degrees the chip's heat-up during operation. And expansion of its temperature range in the low-temperature direction is achieved on account of a corresponding increase in the excitation current in the bubble microcircuit's coils. Information concerning the temperature in the cassette arrives from a heat-sensing integrated circuit placed in it.

The system has two kinds of input/output ports--an RS-232C serial interface, and a parallel byte port that interfaces with microprocessing systems.

The single-card PBM80S module having a capacity of 64K bytes based on 64K-bit bubble microcircuits, and the two-card PBM80M module having a capacity of 2M bits based on 256K-bit microcircuits are offered by Plessey Microsystems, Ltd. Bubble microcircuits are placed on one card of the PBM80M module, and

controller circuits on the other. Both the PBM80S and PBM80M modules are connected directly to an Intel Multibus standard bus. The data transfer rate of the PMB80S module is 100K bytes/s [26].

The combined use of from 1 to 16 Rockwell International RLM658 modules supplemented with an RCM650 programmable control module makes it possible to produce a complete memory system having a capacity of from 128 to 2M bytes [8].

The RLM658 line module contains four RBM256 bubble microcircuits forming a storage array having a capacity of greater than 1M bits and a 256K-bit by 4-bit organization.

The RCM650 control module is software-compatible with the 6502 microprocessor and the Sistema [System] 65 and 6800 design systems.

Domestic developments of bubble memory systems are characterized at present by memory capacities of from 1 [as published] to 16M bits [7, 10, 27].

Testing of Bubble Microcircuits

A distinctive feature of bubble devices is the fact that these devices, unlike tape and disk storages, as a rule do not require outlays for preventive maintenance in the process of their use. However, in the production process manufacturers encounter an entire series of problems associated with the determination of the actual parameters of bubble microcircuits and devices. Specially developed testing equipment is required that is able to synchronize and match in terms of magnitude the magnetic fields that control the movement of bubbles.

Too strong a bias field can cause the destruction of bubbles, and too weak a field the spontaneous generation of excess bubbles or the merging of individual bubbles into domains of the strip type. At elevated temperatures a too strong moving field can also cause the spontaneous generation of bubbles. In either case, there is a loss of information.

Each function implemented by a bubble chip has its own working range. For example, the working range is multidimensional for registers, i.e., it is determined by the amplitude, phase and duration of the current pulse supplied to the excitation circuit, as well as by the strength of the moving field and bias field. Besides, the very architecture of bubble microcircuits, characterized by a high packaging density for shift registers and high synchronization frequencies reaching 1 MHz, requires considerable time for the checking of parameters--on the order of several minutes per microcircuit [29].

The Megatest Corp. Megatest Q-1017 product, called a testing unit for bubble memories and designed for use both in a research laboratory and in manufacturing [10, 29], can serve as an example of a testing system. The testing unit consists of three main parts: a type PDP11/V03 central processing unit with a double floppy disk drive, a rack with analog and digital equipment, and a testing head for connection to bubble devices that are placed in a case or on a board.

Up to eight bubble microcircuits can be connected to the check-out system. The system's software is written in an extended version of PASCAL.

Other automatic testing systems such as Xincom-II, Adate-1450/1475 and Bats-II [10] have also been developed. They are all designed for production or laboratory testing of bubble microcircuits and are similar with respect to characteristics and functional capabilities.

Conclusions

The present state of the art of bubble memories is characterized by bubble devices that have been developed that have a capacity of from 0.25 [as published] to 4M bits and by memory systems based on them having a capacity of from 1M bytes to 16M bytes.

An increase in the capacity of bubble microcircuits to 64M bits can be expected with a change to new technologies for bubble devices.

In spite of the complexity of their manufacture and the complexity of their control electronics, bubble memories will be irreplaceable in the immediate future in military, space and aviation equipment, as well as where highly reliable memory systems are required that work under conditions of external influences and that preserve data when the power is cut off.

Bibliography

1. Posa, D. "National Discontinues Production of Bubble Memories." ELEKTRONIKA, No 18, 1981, pp 7-8.
2. Smith, K. "Plessey Microsystems, Ltd., Continues Involvement With Bubble Memories." Ibid., No 13, 1981, pp 13-14.
3. Manuel, T. "Mass Storage Device." Ibid., No 10, 1986, pp 37-49.
4. Gallakher, R. T. "Fourfold Increase in Capacity of Bubble Memories." Ibid., No 24, 1984, pp 12-13.
5. Nadeyeva, M. B. "New Bubble Memory From Intel." EKSPRESS-INFORMATSIYA. SER. INOSTRANNAYA TEKHNIKA I EKONOMIKA SREDSTV SVYAZI, No 24, 1986, p 5.
6. Parinov, Ye. P. "Element Base of Bubble Memories Using Permalloy Control Coatings." ZARUBEZH. RADIOELEKTRONIKA, No 3, 1983, pp 68-82.
7. Padyukov, L. N. and Rayev, V. K. "Bubble Memories Having Serial Storage Structures." MIKROELEKTRONIKA I POLUPROVODNIKOVYYE PRIBORY. Moscow, Radio i svyaz, No 8, 1984, pp 63-85.
8. Meviti, U. K. "Use of Bubble Memories in Mass Storage." ELEKTRONIKA, No 7, 1979, pp 36-43.

9. George, P. K. and Reyling, Dzh., Jr., "Bubble Memories--State of the Art and Development Prospects," Ibid., No 16, 1979, pp 23-36.
10. Balbashov, A. M., Lisovskiy, F. V., Rayev, V. K. et al. "Elementy i ustroystva na tsilindricheskikh magnitnykh domenakh: Spravochnik" [Bubble Elements and Devices: Handbook], Moscow, Radio i svyaz, 1987. 488 pages.
11. Lomov, L. S., Nurmukhamedov, G. N., Spivakov, D. D. and Chirkin, G. N. "K1605RTs1 Magnetic Integrated Circuit," ELEKTRONNAYA PROMYSHLENNOST, No 4, 1983, pp 16-18.
12. Agafonov, V. A., Garba, L. S., Yekimov, G. N. et al. "Series K1602RTs2 and K1602RTs3 LSI Bubble Memories," Ibid., No 9, 1986, pp 27-30.
13. Sudzuki, R. "Achievements in Area of Bubble Memories," TIIER, Vol 74, No 11, 1986, pp 126-134.
14. Waller, L. "Has the Bubble Memory's Time Come?" ELEKTRONIKA, No 7, 1979, pp 69-71.
15. Nikitov, V. A. "Bubble Devices Having Adjoining-Disk Movement Structures," ZARUBEZHNAЯ RADIOELEKTRONIKA, No 1, 1983, pp 50-70.
16. Waller, L. "Work on Development of 4M-Bit Bubble Memories," ELEKTRONIKA, No 12, 1980, pp 5-7.
17. Posa, D. "Experimental 11.5M-Bit Bubble Memories," Ibid., No 16, 1979, pp 7-9.
18. Manuel, T. "Superhigh-Information-Capacity Memory Devices Based on Bubble Technology," Ibid., No 21, 1985, pp 17-19.
19. Lebose, B. "Start of Production of Memory Boards With Bubble Memories," Ibid., No 26, 1980, pp 10-11.
20. Keypis, R. P. "Start of Industrial Introduction of Bubble Memories," Ibid., No 6, pp 6-7.
21. Brayson, D., Klover, D. and Li, D. "Set of LSI Circuits for Combined Use With Bubble Memories," Ibid., No 9, 1979, pp 23-32.
22. Lafra, R. and Reyling, D. "Controller for Controlling Bubble Memory Systems," Ibid., No 3, 1981, pp 45-52.
23. Shmat, V. K. "Special-Purpose LSI Control Circuits for LSI Bubble Memories," MIKROPROTSESSORNYYE SREDSTVA I SISTEMY, No 6, 1987, pp 3-9.
24. Leboss, B. "Bubble Cassette Memory Operating at Minimum Temperature of -20 °C," ELEKTRONIKA, No 5, 1981, pp 91-92.

25. Beri, L., Lyurta, R. and Reyling, D. "Bubble Cassette Memory With Temperature Compensation." Ibid., No 9, pp 54-58.
26. Smith, K. "Two Boards With Bubble Memories That Can Be Connected Directly to Memory Systems." Ibid., No 19, 1979, pp 13-14.
27. Kuznetsov, S. O., Prokhorov, N. L. and Rayev, V. K. "'Electronic Disk' Utilizing Bubbles." MIKROPROTSESSORNYYE SREDSTVA I SISTEMY, No 4, 1985, pp 11-17.
28. Bisset, S., Bistou, S. and Chen, T. "Features of Testing and Checking of Parameters of Bubble Memories." ELEKTRONIKA, No 10, 1979, pp 45-53.
29. Braunstein, R. "Development of Testing Unit for Bubble Memories." Ibid., No 9, pp 93-94.

COPYRIGHT: Izdatelstvo "Naukova Dumka" "Upravlyayushchiye Sistemy i Mashiny".
1989

HARDWARE

UDC 681.322

Multiprocessor Computing System for Simulation of Radio Systems

907G0030B Kiev UPRAVLYAYUSHCHIYE SISTEMY I MASHINY in Russian No 5, Sep-Oct 89
(manuscript received 28 Aug 88) pp 87-93

[Article by Ye. V. Voronov, A. A. Grigoryev, A. L. Larin and G. I. Donov]

[Text] Present-day complex radio systems have a two-level hierarchical structure. Hardware for the shaping, reception and analog preprocessing of radio signals belongs to the lower, physical level. Algorithms for controlling the hardware of the physical level and for processing streams of digital data arriving from it, making it possible to solve one or another set of application problems of a communications, radar or radio navigation character, belong to the upper, protocol level.

The structure of the algorithms of the protocol level is determined to a considerable extent by the statistical and other characteristics of the streams of data formed at the physical level. The study of these characteristics is an extremely difficult task. It is as a rule not possible to obtain a complete theoretical description of the properties of data streams at the output of physical equipment. And attempts to study these properties by methods of direct digital simulation encounter a lack of sufficient computing capacities. Because of this, the simulation method has become a most important tool for the development and study of the characteristics of complex radio engineering systems.

The method is based on the hardware simulation of data streams formed by the hardware of the physical level by means of what is termed a simulation model of the radio system containing a set of simulators of received signals and a hardware-implemented model of the physical part of the receiving equipment. The simulation model makes it possible to study the algorithms of the protocol level by using data streams that are maximally close in terms of their properties to real streams.

The use of simulation involves the use of computer facilities for controlling the signal simulators and for implementing the data stream processing algorithms. The overall process of servicing a simulation experiment naturally breaks down into a number of computational processes that take place concurrently, that as a rule are closely associated with some structural elements or others of the physical simulation model. The

integration into a single simulation model of structural elements functioning concurrently assumes the interaction of servicing processes for the purpose of exchanging data or of intersynchronization.

An example of the division of a servicing process into concurrent processes is presented in fig 1. A model of the physical part of the receiving equipment implements the concurrent processing of two radio signals, S_1 and S_2 , shaped by simulators IM_1 and IM_2 . Computational processes P_1 and P_2 directly linked to the simulators implement the control of their physical equipment. Process P_3 implements the algorithm for processing the stream of digital data shaped by the receiving equipment, and P_4 plays the role of a session scheduler that furnishes processes P_1 and P_2 with the required system information, including, for example, data concerning the relative dynamics of entities of the system being simulated. Process P_5 provides for total control of the simulation experiment, by implementing the start of processes P_1 to P_4 , monitoring of the course of the simulation session, and the acquisition and processing of experimental data. The most important links between concurrent servicing processes are indicated by arrows in fig 1.

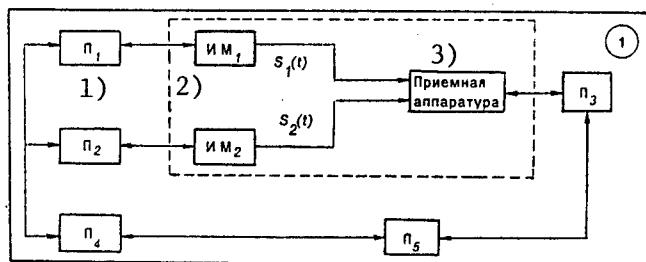


Figure 1. Processes for Servicing Simulation Experiment

Key:

1. P_1	3. Receiving equipment
2. IM_1	

The specific characteristics of the simulation experiment place certain requirements on the architecture of the servicing computing systems. Simulation tasks are satisfied to the greatest extent by multiprocessor computing systems that consist of several relatively independent processor modules furnished with data links and time synchronization and intersynchronization facilities.

A computing system for simulation designed on the basis of KM1810VM86 16-bit single-chip microprocessors [1, 2] is described in the present paper. Process P_5 is implemented on the basis of the DVK-2 interactive computing system, which not only provides for total control of the simulation experiment, but is also used at the preliminary stages as a means of preparing the software, of hardware debugging and of testing the equipment of local processor modules.

System Architecture

The multiprocessor computing system (fig 2) can include up to 14 local processor modules. The nucleus of a local module is a central processing unit constructed from a KM1810VM86 MP [microprocessor] linked to an individual local memory unit (BLP) by means of a local bus, LBUS. Each processor module can be used for servicing a certain structural element of the simulation model. The physical equipment of a servicing element is linked to a local module link via a certain set of input/output registers and can issue service request signals to the local interrupt system.

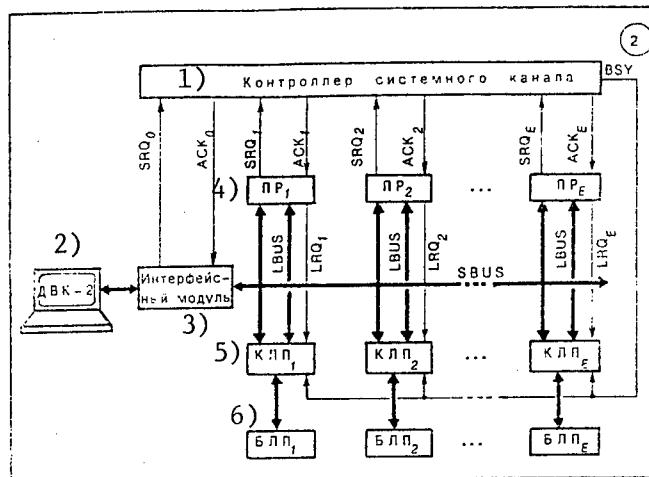


Figure 2. System Architecture

Key:

1. System bus controller	4. PR _j [processor]
2. DVK-2	5. KLP _j [local memory controller]
3. Interface module	6. BLP _j [local memory]

Local processor modules are united into a multiprocessor system by means of a system bus, SBUS, which makes it possible for any processor, PR_j, to access an arbitrary local memory, BLP_k. Local processors use the system bus in the fixed-priority time-sharing mode. Processor PR₁ has the highest bus access priority, and PR_E the lowest. The system bus controller (KSK) exercises bus access arbitration, receives requests from processors (SRQ) and outputs reply signals (ACK) acknowledging the access right, and a system bus busy signal (BSY).

The local memory controller (KLP) registers access requests to local memories arriving through local and system buses and organizes the servicing of these requests. The arbitration of requests at the KLP level is organized according to the cyclic priority principle, i.e., after a local request has been serviced preference is assigned to a system request, and vice-versa.

The structure of the system's system and local buses is similar to the structure of the KM1810 microprocessor bus [1] and differs only in the presence of separate address and data buses. The demultiplexing of the MP's address-data bus is performed inside the processing unit. The make-up of the

system bus, SBUS, (fig 3) include a 20-bit address bus, SA 0 to 19, with an additional byte operation control signal, SBNE; a 16-bit data bus, SD 0 to 15; buffer control signals, SDE and S-OP/IP; read/write control signals, SR/SW; and a reply signal for acknowledgement from the passive end, SRDY. The local bus differs only in the smaller capacity of the address bus, LA 0 to 15, by the presence of two separate inputs for acknowledgment signals from the local memory, LRDY, and local input/output devices, I/O-RDY, and the presence of an additional input/output mode control signal, M/I/O.

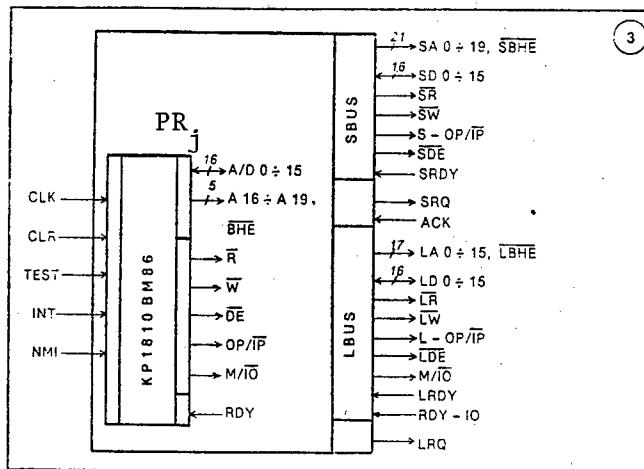


Figure 3. Local Processor Module Interface

The KM1810 MP's memory access request is physically a dynamic procedure (termed bus cycle) in which it is possible to distinguish an addressing phase, a dynamic initial phase, a static acknowledgement waiting phase and a dynamic termination phase (fig 4). In the addressing phase the MP outputs a complete physical address in line A/D 0 to 15 and A16 to A19 and accompanies it with an address strobe, STB. During the dynamic phases of the start and end of a cycle it activates and turns off control signals DE, R-bar and W-bar. The output and turning off of these signals are performed in a certain order that makes possible the correct functioning of bus buffers and memory elements. The duration of dynamic phases is strictly assigned by the microprocessor. The static access phase begins from the instant control signals are established and ends after the arrival of the acknowledgement signal, RDY, from the addressed passive sharer. The duration of the static phase is not limited and is determined at the discretion of the passive end.

Before arriving in the memory of an individual local memory unit, the access request formed by the MP travels a certain route in the bus network of the multiprocessor system.

First, requests are divided at the processor level into local and system. Requests are separated in the addressing phase. Then local requests enter the local bus and system requests are localized within the central processing unit for the arbitration period. They are output to the system bus only upon the ACK signal for access right acknowledgement from the KSK [system bus

controller]. On the local memory controller level, requests go through a second stage of arbitration with respect to access right to the memory.

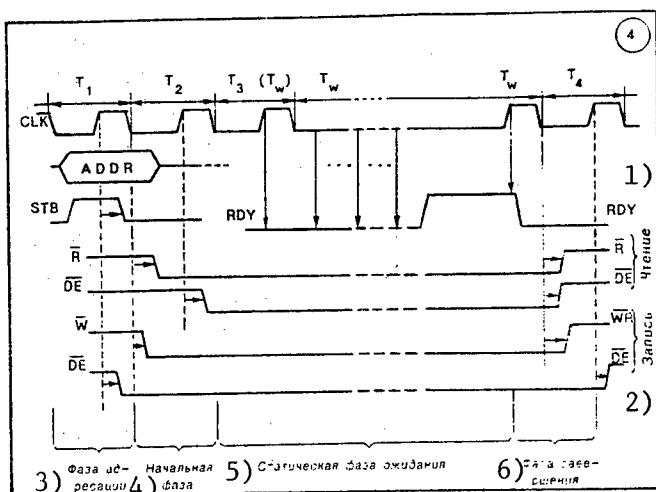


Figure 4. Structures of Read/Write Cycles (T_w --Wait Cycles)

Key:

1. Read	4. Initial phase
2. Write	5. Static waiting phase
3. Addressing phase	6. Termination phase

A certain amount of time, which is spent on arbitration and line switching, is required for a request to travel the route designated for it. The request remains in the static waiting phase during this whole time. RDY acknowledgement signal propagation circuits are switched concurrently with line switching that makes possible the serial advancing of a request from the MP to a memory, so that the circuit for transmitting the RDY reply signal from this memory to the RDY input of the MP is found to be prepared by the instant a request physically enters the line of one of the memories. Upon receipt of the acknowledgement signal, the MP introduces the access termination phase, whose passing causes the removal of requests LRQ and SRQ. As a result, the connections enabling a request to travel along the route are broken.

The principle, on which the organization of the bus network is based, of arbitration and switching during the static waiting phase creates certain complications in the organization of memory access. The fact is that, as a rule, a request enters a memory already in the static phase, i.e., after the termination of the dynamic start-of-access procedure formed by the MP. This means that the time sequence of instants when control signals are "turned on", which is required for the proper functioning of memory elements, must be formed artificially.

In the system under discussion, the task of converting static requests into dynamic memory access procedures is entrusted to local memory controllers, which, having begun the servicing of a local or system request, make it

possible to switch control signals on in the necessary time sequence. The time sequence for turning control signals off is formed by the microprocessor itself during the cycle termination phase.

A request's route in the system's bus network is assigned by a 4-bit segment number code output by the MP in lines A16 to A19 in the addressing phase. The MP's address space is divided into 16 segments, SEG0 to SEGF, 64K bytes in size [3]. Calls to segments SEG0 and SEGF are sent by the processor into the local bus. A local request signal, LRQ, is formed. But calls to all remaining segments of the address space cause the formation of a system request signal, SRQ. These requests are output to the system bus together with the segment number code upon the acknowledgement signal, ACK, from the system bus controller (KSK). Simultaneously with the ACK signal, the KSK generates the system bus busy signal, BSY, which informs all local memory units of the presence of a system request in SBUS lines.

Segments SEG0 and SEGF are found to be "linked" in the local processor's address space, in the sense that calls to them according to identical relative addresses are converted into identical local requests. This makes it possible to use a local bus both in initial start-up of the MP through the CLR input (calls to segment SEGF) and in the retrieval of interrupt vectors (calls to segment SEG0). The MP's input/output cycles are also sent into the local bus. However, the LRQ signal is not generated with this. Segments SEG1 to SEGE of the MP's address space are used for access to "foreign" local memory units via the system bus.

A local memory unit is mapped at once onto three segments of the address space. It senses local requests addressed to segments SEG0 and SEGF and, besides, with an active level of signal BSY, system requests addressed to segment SEG1.

If the route is transparent (there are no delays for arbitration), the bus network of the system under discussion makes it possible for local memory calls to pass through during four timing cycles, CLK. Thus, in the absence of system requests the speed in local buses is determined only by the MP's speed. When a conflict with a system request originates, the duration of a local access cycle increases. The time it takes for a system call to pass through with a transparent route is five cycles. The inclusion of one additional wait cycle is due to delays for arbitration at the system bus controller level.

Synchronization Facilities

The facilities offered by the multiprocessor system for intersynchronization and time synchronization can be divided into software, firmware and hardware facilities.

At the software level the intersynchronization of processes is made possible by the use of communication bytes placed in local memory modules. The contents of these bytes determine the nature of the current local process and can be set by other processes via the system bus.

An example of the software organization of what is termed the point of encounter of process P_j with process P_k is shown in fig. 5. In this figure process P_k , having terminated the execution of program module PROG, reports to process P_j concerning this. For this purpose, it increases by one the contents of communication byte F located in the local memory for process P_j . Process P_j waits for a message from process P_k , by cyclically polling the communication byte. It continues to run its own program after the receipt of the message. The use of communication bytes makes possible the software organization of quite different protocols for interaction between concurrent processes.

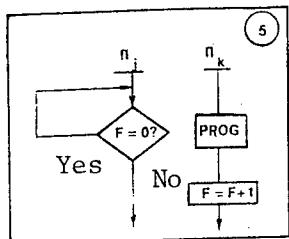


Figure 5. Principle of Synchronization at Software Level

Synchronization at the software level creates an extra load on the system bus that can be eliminated in principle by removing the communication bytes beyond the memory's bounds and organizing direct wire connections between them. The system's firmware facilities can be regarded as one possible implementation of this approach.

Each processor module is furnished with a 16-bit control register, CR, and status register, SR, (fig 6) that are located in local address input/output spaces. The output of the k-th bit of register CR_j is electrically connected to the input of the j-th bit of register SR_k . The individual bits of CR registers are used for transmitting messages to processes having the corresponding numbers, and the bits of SR registers for receiving messages from other processes. Thus, the bits of CR and SR registers play the role of "communication bits" used for the organization of protocols for the interaction of concurrent processes. The outputs of the zero bits of CR registers are equipped to transmit messages to the DVK-2 central microcomputer.

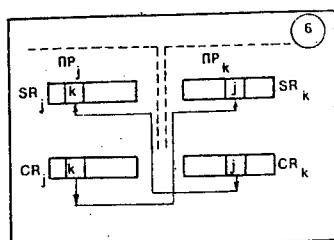


Figure 6. Status (SR) and Control (CR) Registers

In addition to intersynchronization, simulation experiment servicing processes as a rule require referencing to a certain single time scale formed by the system timer. The time synchronization of processes is made possible by facilities of the hardware level on account of the supplying of timer signals to the MP's TEST inputs.

DVK Interface

The DVK communicates with the multiprocessor system via an interface module (cf. fig 2) that makes it possible to interface the DVK's system bus to the system's system bus, SBUS, and offers certain facilities for controlling local processors and checking their status. A software model of the interface is shown in fig 7.

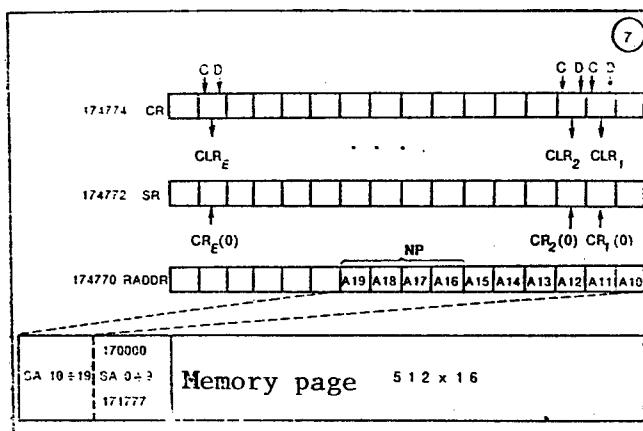


Figure 7. Software Model of DVK Interface (NR--Processor Number Register)

The interface provides the DVK access to any of 14 local memory units via the SBUS bus. One page of local memory 512 16-bit words in size is open for access from the DVK at any given instant. This page is mapped onto addresses 170000 to 171777 in the DVK's address space. The belonging of an open page to one local memory unit and its place inside this unit are assigned by a code established in the interface's address register, RADDR. The module converts the DVK's access to a memory page into the corresponding access to the system bus. Here the 10 least-significant bits of the 20-bit system address, SA, are assigned by the relative access address within the page, and the 10 most-significant bits are assigned by the contents of RADDR.

The DVK interface uses the system's system bus together with local processor modules in the fixed-priority time-sharing mode and has the highest priority.

The interface module contains 16-bit control, CR, and status, SR, registers. The inputs of the status register are connected to the outputs of the zero bits of the control registers, CR_j, of the local processor modules and are used for receiving messages from them. The control register makes it possible

for the DVK to execute a programmed start and stop of local processors. Its bit outputs are connected to the MP's start-stop CLR inputs.

The allocation of bits 0 and 15 of registers CR and SR is not stipulated. Each bit of the control register has an independent synchronous write bus represented by a data line, D, and write strobe line, C, and can be interfaced to the local bus of one of the processor modules as an independent one-bit output register. As a result of this interface, a local module gains the ability to control the start-stop of another local module without the DVK's assistance.

System's Monitor System

The multiprocessor computing system's monitor system (MS), developed on the basis of the DVK RT-11 operating system (OS), is a software package used in the preparation, debugging and performance of simulation experiments. The MS contains facilities for software preparation and debugging, facilities for testing the system's hardware, facilities for checking out the simulation model hardware's interfaces to local buses, and facilities for organizing the DVK control computer's interface to the system during the course of an experiment.

The MS offers the capability of developing programs in KR1801 (DVK) and KM1810 (local processors) MP codes by using assemblers. The source text files of programs are created and edited by means of the RT-11 OS K52 editor. A standard set of OS facilities is used for writing programs in KR1801 MP codes. Programs in KM1810 local processor codes are written by means of the AS8086 cross assembler, which converts the type .S input text files in KM1810 MP assembly language into type .REL output object files. Files of the .REL type are text files and contain the operation codes of instructions in a hexadecimal byte representation with an indication of hexadecimal absolute loading addresses. The loading of these files for execution presupposes the conversion of hexadecimal text data into binary codes. This conversion is performed by a special loading routine, LOADER, included in the MS.

The MS's structure is shown in fig 8. Its nucleus is the CONSOL console program that provides the operator a set of facilities for communicating with the system through instructions entered from the DVK's keyboard. The console program is started from the RT-11 OS as a user program and functions together with local monitors, LMON, resident in processor module memory units.

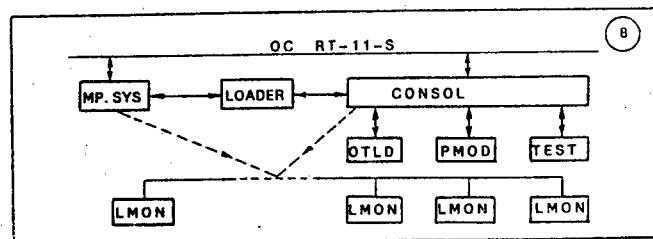


Figure 8. Structure of Monitor System

Other MS program modules (file loader of the .REL LOADER type, TEST test package, OTLD debugger and PMOD periodic execution mode control module) are closely linked to the console program and are started from it through the appropriate commands. A relatively independent unit of the MS is the MP.SYS multiprocessor system driver, which is used both together with the console program while the loader is operating and independently as a driver of one of the devices serviced by the RT-11 operating system.

CONSOL Instruction Set

In starting from the RT-11 operating system, the CONSOL program loads local monitors, LMON, into all local memory units and switches to the mode of communication with the operator. When necessary, LMON loading can be performed repeatedly for each processor module individually.

The CONSOL instruction set makes it possible to perform the following operations:

To open for operations any local module, by setting its number in the processor number register, NP.

To open and modify the contents of the words and bytes of the local memory unit (BLP) of the current (open for operations) module. Access to a memory element begins with the input of its local address, which can be given in the octal or hexadecimal formats. The contents of a memory element can be "opened" in the octal, byte octal, hexadecimal and binary formats.

To open and modify the contents of MP internal registers--general-purpose registers--as well as their lowest- and highest-order bytes, and of pointer registers, segment registers, a program counter and flag register. In operations with registers, their contents are assessed and written to individual LMON memory locations. When the user program is started, LMON makes it possible to load the contents of these locations into MP internal registers.

To open the contents of an interface module status register, as well as to open and modify the contents of a control register.

To perform operations of a single exchange of data in input/output with registers of external devices connected to the local bus of the current processor module. A register is accessed by specifying its address. Eight- and 16-bit registers can take part in exchange operations. Direct access of the DVK processor to local buses of processor modules is not possible. Therefore, local registers of external devices are accessed indirectly by means of a local monitor, LMON. The DVK transmits through the system bus to the LMON an access requisition containing the register's address, its capacity, the exchange direction and the data to be output and starts the processor module. LMON implements the required exchange operation and reports concerning its completion to the DVK via control and status registers, having first placed in memory the data input from the register.

To start the routines, included in the TEST package, for testing the system's hardware.

To call the loader, LOADER, of files of the .REL type with programs in KM1810 MP codes created by the AS8086 cross assembler.

To call the OTLD software debugger.

To call the PMOD module, which organizes the periodic execution of short test routines for the observation on an oscillograph of processes taking place.

Debugging Mode

The OTLD debugger program makes it possible to debug programs in the serial and parallel execution modes. The mode is selected through the appropriate CONSOL commands.

One program resident in the memory of the current processor module is debugged with serial execution. This program can be started in the step-by-step execution mode (command G) or in the mode of continuous execution with stops at checkpoints (KTs) (command R). The step-by-step mode is organized on the basis of T-bit interrupts. Checkpoints are set and removed by CONSOL facilities. Checkpoints are introduced by replacing the first byte of the stop address operation code by the code 314 of command interrupt INT3. When the checkpoint is removed the original operation code is automatically restored.

The OTLD program, local monitor (LMON) and program to be debugged (PROG) take part in the organization of debugging modes. At the G or R command the OTLD program sets or clears, respectively, the T-bit in the MP's flag register and starts the processor module after having turned off the CLR signal. Then it goes into the supervisor mode.

After the CLR signal has been turned off, the LMON monitor loads the internal registers of the MP and transfers control to the user program (PROG) through the RTI command, which enables loading from the stack of the flag register, IF, with a set or cleared T-bit and of a complete start address (registers CS and IP). If the T-bit has been set (the step-by-step mode), then after the execution of one instruction of the program being debugged a T-bit interrupt occurs and control is returned to LMON. Otherwise the user program is run continuously until one of the INT3 checkpoints is reached, after which control is also transferred to LMON.

Having received control through one of the debugging interrupts, LMON saves in memory the contents of the MP's internal registers, sends via the CR register a message concerning the halt of the program being debugged, and is stopped through the HLT command. In reply to the message regarding the halt, the debugger stops the MP through the CLR input, outputs diagnostic data to the display, and changes to the mode of communication with the operator.

If a halt of the user program does not occur, then OTLD is "suspended" in the supervisor mode. The possibility is provided of a forced exit from this mode to the mode of communication with the operator.

In the parallel debugging mode, a list of the numbers of the modules taking part in debugging and a list of the numbers of the modules that are to be started in the step-by-step execution mode are transferred to the OTLD program. These lists are defined by CONSOL facilities by setting ones in the appropriate bits of the parallel-process register, PP, and T-bit register, TB. The parallel starting of processes tagged by ones in register PP is executed through the R command. With this, processes tagged in register TB are started in the step-by-step mode. The debugger remains in the supervisor mode until all started processes have been halted.

Periodic Execution

The PMOD module organizes the periodic execution of testing routines stored in the DVK's memory or in local memory units of processor modules.

Testing routines in the DVK's memory that are used in finding faults in an interface module are created by special facilities of CONSOL. When they are started in the periodic mode, the PMOD program provides for the generation of a sync signal for triggering an oscilloscope at the start of each execution cycle.

Testing routines are loaded into local memories from MP initial start address 177760. The PMOD module makes possible the "loop running" of these routines by the software generation of a periodic signal in the CLR inputs of the microprocessors. Local testing routines can be started for periodic execution in the serial or parallel execution modes. In the latter instance the make-up of processes to be started concurrently is determined by the contents of the parallel-process register, PP.

The availability of periodic modes renders an invaluable service in finding faults and debugging interfaces between simulation model hardware and local buses.

Test Package

The testing routines included in the test package make it possible to check out communication between the DVK and interface module registers, communication between the DVK and local memories via the system bus, communication between local processor modules and "their own" memories via the local bus, and communication between local processors and "foreign" memories via the system bus. The CONSOL program's commands make it possible both to call individual tests and to organize various cyclic chains of tests. When failures are detected, all test programs output diagnostic data and automatically generate in the DVK's memory or local memory testing routines that can be started in the periodic execution mode for the purpose of revealing the reason for the failure.

System Driver

The MP.SYS driver makes it possible to include the multiprocessor system among the number of peripheral devices serviced by the RT-11 operating system. This makes it possible to perform certain operations with the system at the command and program levels by the operating system's facilities.

The MP.SYS driver defines an MP (multiprocessor system) device as a device having a non-file structure that permits read and write calls as well as .SPFUN program requests for the performance of special functions. The driver services one of 14 local modules at any given instant. The number of the module to be serviced is a parameter of the driver and can be set through the SET command or .SPFUN program request.

One block of arbitrary size can be opened for input/output in the local memory of the module being serviced. The block's attributes (its start address, length in bytes and type) are set through SET parameter assignment commands or through an .SPFUN special program request. A block can be declared a binary or text block with respect to type. Binary blocks are used for the input/output of data in binary format.

In operations with blocks of the text type, the driver converts the data to be written from hexadecimal text format (.REL file format) to binary, and data to be read, from binary format to text. The codes contained in files of the .REL type are loaded according to the addresses indicated in these files themselves. When text blocks are read, the contents of the memory area defined by the attributes are converted into .REL type test file format, which can be written to disk or printed. In addition to data exchange operations, the driver services special program access requests to interface module CR and SR registers.

The set of functions offered by the driver simplifies the use of the processor for the overall control of a simulation experiment. The driver makes it possible to develop control programs by using standard assembler macro facilities for organizing communication with the system.

Bibliography

1. Kobylnskiy, A. V., Moskalevskiy, A. I. and Temchenko, V. A. "KM1810VM86 Single-Chip High-Performance 16-Bit Microprocessor." MIKROPROTSESSORNYYE SREDSTVA I SISTEMY, No 1, 1986, pp 28-33.
2. Kobylnskiy, A. V., Bereza, N. G., Sabadash, N. G. et al. "KM1810VM86 Microprocessor Instruction Set." Ibid., No 2, pp 3-9.
3. Grigoryev, V. L. "Programmirovaniye odnokristalnykh mikroprotsessorov" [Programming of Single-Chip Microprocessors]. Moscow, Energoatomizdat, 1987, 286 pages.

COPYRIGHT: Izdatelstvo "Naukova Dumka" "Upravlyayushchiye Sistemy i Mashiny", 1989

APPLICATIONS

UDC 658.512.2:681.3

A Parallel Bolder Algorithm for Rotation Operations and its Use in Computer Graphics

907G0097b Kiev UPRAVLYAYUSHCHIYE SISTEMY I MASHINY in Russian No 1, Jan 90 pp 106-109 (manuscript received 04 Jan 87)

[Article by Ye. I. Artamonov, Sh.-M. A. Ismailov, O. G. Kokayev, V. M. Khachumov]

[Text]

Algorithms of two-dimensional rotation are at the basis of many important applications of computer graphics. Above all, one should separate the generation of sequential frames of images, when each subsequent frame differs from the previous one by a rotation by some degrees, which creates the appearance of smooth rotation of the object [1]. The use of traditional methods of determining the coordinates of the rotated image in this case requires one to calculate trigonometric functions, and to carry out multiplication operations and floating point additions; this makes it difficult to implement the rotation in real time.

The time can be reduced by using special algorithms which do not contain labor-intensive operations and which can be easily implemented by the hardware. The algorithm of D. Bolder is interesting [2]. In this algorithm, the smooth rotation of a vector around the origin of the coordinates at an angle ϕ is replaced by a sequence of sign changing rotations about an angle ϕ_i ($i = 1, \dots, n$).

Transformation of the system of coordinates is described by the expression [3]:

$$\begin{bmatrix} K & X' \\ K & Y' \end{bmatrix} = \begin{bmatrix} n \\ \prod_{i=1}^n l_i \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}, \quad (1)$$

where

$$l_i = \begin{bmatrix} \cos \varepsilon_i \varphi_i & \sin \varepsilon_i \varphi_i \\ -\sin \varepsilon_i \varphi_i & \cos \varepsilon_i \varphi_i \end{bmatrix}, \quad (2)$$

and in the binary system:

$$\begin{aligned} \varphi_{i+1} &= \varphi_i - \varepsilon_i \operatorname{arctg} 2^{-i}, \\ \operatorname{sign} \varepsilon_i &= \operatorname{sign} \varphi_i, \quad \varphi_r = \varphi. \end{aligned} \quad (3)$$

Here $[X, Y]$, $[X', Y']$ are, respectively, the old and new coordinates of the end of the vector which is rotated counterclockwise; n is the number of rotations which

defines the maximum size of binary numbers which give the coordinates; ϵ_i ($\epsilon_i \in \{-1, 1\}$) are operators which define the direction of the i^{th} angle of rotation; K is the coefficient of elongation of the vector for $n = 10$ $K \approx 1.65$.

The traditional approach to implementing the system (1)-(3) is based on the use of the "digit after digit" iteration method [3]. The recursive formulas of this method include operations of multi-bit shift, addition, and subtraction of numbers in fixed point form, which leads to an increase in the speed and a decrease in the hardware costs to carry out the calculations. Various approaches to hardware implementation of recursive formulas presented in [4] as a whole do not change the sequential character of the calculations and make it possible to obtain ordered coordinates of points of the rotated image after carrying out $i \geq n/2$ iterations. An exception is the principle of conveyer shifting of iteration cycles; however, this requires the use of significantly more complicated equipment.

This work proposes a fast rotation algorithm based on the parallel organization of the computing process, using expanded recursive formulas of the "digit after digit" method.

The algorithm which has been developed includes the following basic stages: placement of expressions (1)-(2) in parallel form, defining all operators ϵ_i according to formula (3) for fixed angles of rotation which were chosen ahead of time, calculation of the coordinates of points of the rotated image based on a group summation operation. Let us examine each of these stages in more detail.

For definiteness we will assume that the plane of the screen of a graphics terminal is 512×512 , and we will use ten-bit binary operands to represent the coordinates of the points of the image. Then, considering the acceptable error of calculation is less than the unit of the low bit, the total length of the bit grid is 16.

When the initial Bolder correlations (1) and (2) are parallelized for $n = 10$ we obtain, respectively,

$$KX' = B2^0 - \sum_{k=1}^{10} \epsilon_k A2^{-k} - \sum_{k=1}^4 \sum_{l=k+1}^{10-k} \epsilon_k \epsilon_l B2^{-(k+l)} + \sum_{k=1}^{[(9-k)/2]} \sum_{l=k+1}^{10-(k+l)} \sum_{m=l+1} \epsilon_k \epsilon_l \epsilon_m A2^{-(k+l+m)}, \quad (4)$$

$$KY' = A2^0 + \sum_{k=1}^{10} \epsilon_k B2^{-k} - \sum_{k=1}^4 \sum_{l=k+1}^{10-k} \epsilon_k \epsilon_l A2^{-(k+l)} - \sum_{k=1}^{[(9-k)/2]} \sum_{l=k+1}^{10-(k+l)} \sum_{m=l+1} \epsilon_k \epsilon_l \epsilon_m B2^{-(k+l+m)}, \quad (5)$$

where $A = X + \epsilon_0 Y$, $B = Y - \epsilon_0 X$, $[x]$ is the whole part of x .

Further expansion of the resultant expression is not expedient due to a significant increase in the number of terms; moreover, the calculation of A and B may coincide with the stage of preparing ϵ_i ($i = 1$ to 10). To organize the calculations on formulas (4) and (5) one must know the values of all the operators and the corresponding given angle of rotation.

Table 1 shows the operators ϵ_i ($i = 0$ to 10) for fixed angles in the range $0-90^\circ$ in the counterclockwise direction. The values that are found can be conveniently

Table 1.

Угол α (в градусах)	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
0	1	-1	-1	-1	1	-1	1	1	-1	-1	-1
10	1	-1	-1	-1	-1	1	1	-1	-1	1	-1
20	1	-1	1	-1	-1	-1	1	-1	-1	-1	-1
30	1	-1	1	-1	1	1	-1	-1	-1	-1	-1
40	1	-1	1	1	-1	-1	-1	-1	-1	-1	-1
50	1	-1	1	-1	-1	1	1	1	-1	-1	-1
60	1	-1	1	-1	-1	1	1	-1	1	-1	-1
70	1	-1	1	1	-1	-1	1	1	1	-1	-1
80	1	-1	1	-1	1	-1	-1	1	1	-1	-1
90	1	1	1	1	-1	1	-1	-1	1	1	1

a. angle in degrees.

stored in ROM with parallel sampling. We note that to rotate the image in the clockwise direction the signs of the operator must be changed to the opposite sign.

Parallel calculation of expressions (4) and (5) becomes possible due to the original operation of group summation proposed in [5]. Let us give the result of summation of N binary numbers of size m in the following form:

$$Z = S_m S_{m-1} \dots S_1,$$

where S_i is the value of the i^{th} bit of the result.

Since the terms in (4) and (5) are formed by a shift of the same operand A or B by a variable number of steps in one direction, then it is assumed that overflow does not occur, and additional bits are not needed to represent the result. In turn, S_i is defined by the sum of values S'_i of the i^{th} bits of all N terms, representing the i^{th} bit cut-off, considering the transfer of P_{i-1} from the $(i-1)^{th}$ bit cut-off, and is written in the form

$$S'_i = P_{i-1} \oplus \sum_{j=1}^N S'_j \pmod{2},$$

$$P_i = \left[\left(\sum_{j=1}^N S'_j + P_{i-1} \right) / 2 \right], \quad P_0 = 0.$$

Considering the validity of the notation $[(L + M)/2] = [L/2] + [(L \pmod{2} + M)/2]$ for whole L and M , the transfer of P_i may be given in the form of the sum $P_i = P_i^1 + P_i^2$, respectively the basic and supplemental transfers, so

$$P_i^1 = \left[\left(\sum_{j=1}^N S'_j / 2 \right) \right],$$

$$P_i^2 = [(S_i + P_{i-1}^1 + P_{i-1}^2) / 2], \quad P_0^1 = P_0^2 = 0.$$

If in the memory one first writes all possible values of multi-bit transfers of P_i^1 and P_i^2 , one can use them to construct an absolute summing parallel device.

Figure 1 shows the structure of such a device. It contains the following: an input register (IR) to receive the current bit cut-off; associative and information parts

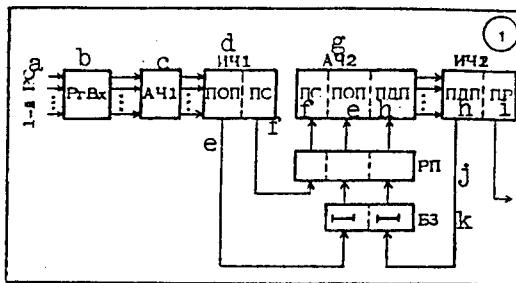


Figure 1. Associative summing device. a. first bit cut-off; b. input register; c. associative part 1; d. information part 1; e. field of main transfer; f. field of sum by modulo two; g. associative part 2; h. field of additional transfer; i. field of the result; j. transfer register; k. delay unit.

(AP1, IP1) of the first associative memory; associative and information parts 2 (AP2, IP2) of the second associative memory, which identify and store the main and additional transfers, the results of summing by modulo two, and the values of S_i ; a transfer register TR; a delay unit DU. The following notations are used: field of main transfer FMT, field of additional transfer FAT, field of the sum by modulo two FS, field of the result FR.

Figure 2 shows, as an example, information written in the associative memory in the addition of five 16-bit operands. The time in cycles to complete the operation of group summing is $T = m + \lceil \log_2 N \rceil$, where $\lceil \cdot \rceil$ indicates the closest, smallest whole. The number of terms in each expression (4), (5) is $N = 42$ for $m = 16$, and the time for their parallel summation is about a factor of 7 smaller than the time to sum the same number of terms in known fast summers for two operands.

The general structure of the rotation device based on Bolder's parallel algorithm is shown in Figure 3. The device contains the following: preliminary summers S1, S2 to form values A and B; registers R1, R2 for the parallel shift and storage of the results of preliminary summing; ROM for the storage and issuing of operators corresponding to the given angle of rotation; combination circuits CC1 and CC2 which form the signs and additional codes of the terms; devices S3 and S4 to carry out the group summation operation.

Table 2 gives the results of modeling of the rotation operation for points with initial coordinates [20, 20]. In addition to the values KY' and KX' obtained on the basis of the parallel Bolder algorithm, we give coordinates Y' and X' calculated from rotation formulas which are traditionally used in computer graphics [1], as well as the values of KY'/K and KX'/K . It is clear from the table that the absolute error of the Bolder transformation after appropriate normalization of the results is within acceptable limits, which provides good quality imaging.

The observed effect of lengthening the position vectors is an attribute of the Bolder algorithm. At present, to counter it, an algorithmic method of compensating for the deformation is used which involves more complicated iteration equations, and, as a result, an increase in the hardware [4]. Another solution to the problem is scaling of the result, since the value of K is known beforehand. Since only the whole part of coordinates X' and Y' are used for imaging, the scaling problem is simplified and may be solved by attaching code transformers to the outputs of the rotation device.

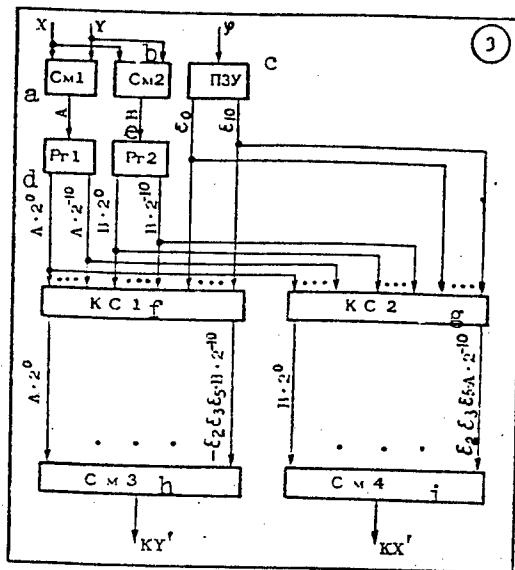
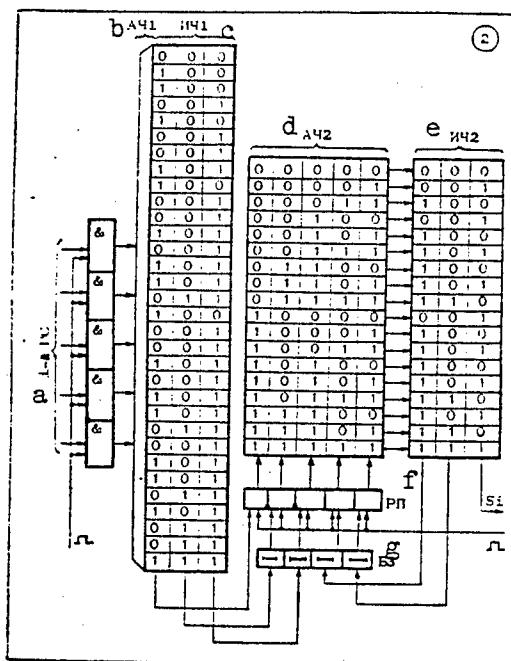


Figure 2 [left]. Example of writing information to the associative memory. a. first bit cut-off; b. associative part 1; c. information part 1; d. associative part 2; e. information part 2; f. transfer register; g. delay unit.

Figure 3 [right]. Structure of the rotation device. a. S1; b. S2; c. ROM; d. R1; e. R2; f. CC1; g. CC2; h. S3; i. S4.

Table 2.

α Угол (в градусах)	KY'	KY'/K	Y'	KX'	KX'/K	X'
0	33,05	20,03	20,00	32,93	19,96	20,00
10	38,20	23,15	23,17	26,68	16,17	16,22
20	42,34	25,66	25,63	19,65	11,91	11,95
30	44,92	27,23	27,32	11,99	7,27	7,32
40	46,41	28,13	28,18	4,10	2,49	2,47

a. Angle in degrees.

Such a transformer could be made with mass-produced ROM microcircuits (for example, two K556RT7 microcircuits). The algorithm which has been developed can be easily modified for the case of rotation around a random point of the screen.

The use of the principles of group summing in the algorithm provides a significant acceleration of the processes of transformation and imaging of graphic information.

BIBLIOGRAPHY

1. Foli, Dzhv; Ven Dem, A. "Osnovy interaktivnoy mashinnoy grafiki" [Foundations of interactive computer graphics] Moscow: Mir 1985 Book 1 368 pp.

2. Bakhtiarov, G. D., Zinyakov, N. L. Bolder's algorithm and its use in computing and digital signal processing ZARUBEZH. RADIOELEKTRONIKA 1983 No 3 pp 3-23.
3. "Mini-EVM. Printsipy postroyeniya i proyektirovaniya" [Minicomputers. Principles of construction and design] A. V. Palagin, V. A. Ivanov, A. F. Kurgayev, V. P. Denisenko. Kiev: Nauk. dumka 1975 175 pp.
4. Baykov, V. E., Smolov, V. B. "Spetsializirovannyye protsessory. Iteratsionnyye algoritmy i struktury" [Specialized processors. Iteration algorithms and structures] Moscow: Radio i svyaz 1985 288 pp.
5. Ismailov, Sh.-M. A. "Assotsiativnyye protsessory parallelnoy obrabotki potokov chislovых dannykh v komplekse sistem schisleniya" [Associative processors for parallel processing of digital data flows in a calculation system complex]. Author's abstract of the Dissertation for Candidate of Technical Sciences. Leningrad: LETI 1985 24 pp.

- END -

10
22161

55

NTIS
ATTN: PROCESS 103
5285 PORT ROYAL RD
SPRINGFIELD, VA

22161

This is a U.S. Government publication. Policies, views, or attitudes of the U.S. Government. Users of this publication may cite FBIS or JPRS provided they do so in a manner clearly identifying them as the secondary source.

Foreign Broadcast Information Service (FBIS) and Joint Publications Research Service (JPRS) publications contain political, economic, military, and sociological news, commentary, and other information, as well as scientific and technical data and reports. All information has been obtained from foreign radio and television broadcasts, news agency transmissions, newspapers, books, and periodicals. Items generally are processed from the first or best available source; it should not be inferred that they have been disseminated only in the medium, in the language, or to the area indicated. Items from foreign language sources are translated; those from English-language sources are transcribed, with personal and place names rendered in accordance with FBIS transliteration style.

Headlines, editorial reports, and material enclosed in brackets [] are supplied by FBIS/JPRS. Processing indicators such as [Text] or [Excerpts] in the first line of each item indicate how the information was processed from the original. Unfamiliar names rendered phonetically are enclosed in parentheses. Words or names preceded by a question mark and enclosed in parentheses were not clear from the original source but have been supplied as appropriate to the context. Other unattributed parenthetical notes within the body of an item originate with the source. Times within items are as given by the source. Passages in boldface or italics are as published.

SUBSCRIPTION/PROCUREMENT INFORMATION

The FBIS DAILY REPORT contains current news and information and is published Monday through Friday in eight volumes: China, East Europe, Soviet Union, East Asia, Near East & South Asia, Sub-Saharan Africa, Latin America, and West Europe. Supplements to the DAILY REPORTS may also be available periodically and will be distributed to regular DAILY REPORT subscribers. JPRS publications, which include approximately 50 regional, worldwide, and topical reports, generally contain less time-sensitive information and are published periodically.

Current DAILY REPORTS and JPRS publications are listed in *Government Reports Announcements* issued semimonthly by the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, Virginia 22161 and the *Monthly Catalog of U.S. Government Publications* issued by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

The public may subscribe to either hardcover or microfiche versions of the DAILY REPORTS and JPRS publications through NTIS at the above address or by calling (703) 487-4630. Subscription rates will be

provided by NTIS upon request. Subscriptions are available outside the United States from NTIS or appointed foreign dealers. New subscribers should expect a 30-day delay in receipt of the first issue.

U.S. Government offices may obtain subscriptions to the DAILY REPORTS or JPRS publications (hardcover or microfiche) at no charge through their sponsoring organizations. For additional information or assistance, call FBIS, (202) 338-6735, or write to P.O. Box 2604, Washington, D.C. 20013. Department of Defense consumers are required to submit requests through appropriate command validation channels to DIA, RTS-2C, Washington, D.C. 20301. (Telephone: (202) 373-3771, Autovon: 243-3771.)

Back issues or single copies of the DAILY REPORTS and JPRS publications are not available. Both the DAILY REPORTS and the JPRS publications are on file for public reference at the Library of Congress and at many Federal Depository Libraries. Reference copies may also be seen at many public and university libraries throughout the United States.